

Tilburg University

Zelf aan het stuur en op maat. Werkt het en waarom wel of niet?

Muffels, Ruud; Blom-Stam, Kirsten; van Wanrooij, Stefan

Publication date:
2020

Document Version
Publisher's PDF, also known as Version of record

[Link to publication in Tilburg University Research Portal](#)

Citation for published version (APA):

Muffels, R., Blom-Stam, K., & van Wanrooij, S. (2020). *Zelf aan het stuur en op maat. Werkt het en waarom wel of niet? Eindverslag Vertrouwensexperiment Apeldoorn/Epe*. Tilburg University/ReflecT/Tranzo.

General rights

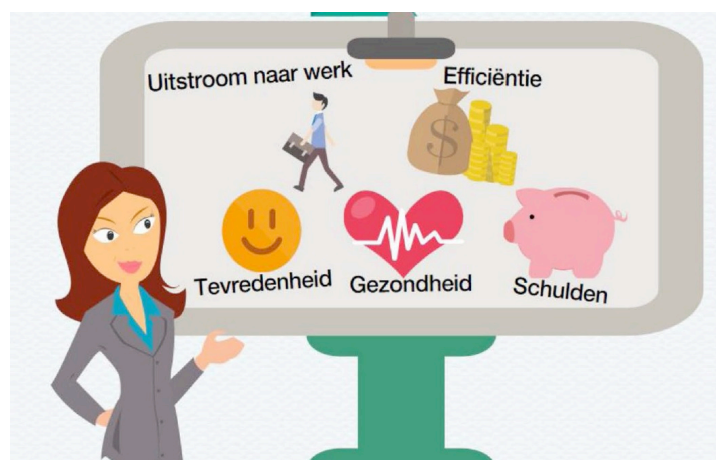
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

**Zelf aan het stuur en op maat.
Werkt het en waarom wel of niet?
Eindverslag Vertrouwensexperiment Apeldoorn/Epe**



*Ruud Muffels (projectleider), Kirsten Blom-Stam (onderzoeker)
& Stefan van Wanrooij (student assistant)*

Tilburg-Apeldoorn/Epe, 8 Mei 2020
© Tilburg University/Reflect

Inhoudsopgave

1. Inleiding (doel concept rapportage; inhoud rapportage, databronnen)	2
1.1 Aanleiding voor het onderzoek; beleidsvragen gemeente	4
1.2 Theoretisch kader	5
1.3 Onderzoeksvragen	14
2. Opzet onderzoek	16
2.1 Experimentgroepen	16
2.2 Procedure van random selectie en toewijzing	17
2.3 Onderzoekspopulatie	19
2.4 Product- en procesevaluatie	22
2.5 Dataverzameling product- en procesevaluatie	23
2.5.1 Productevaluatie: vragenlijsten	23
2.5.2 Administratieve BUS data	25
3 Resultaten	26
3.1 Productevaluatie op basis van uitkomstmaten	36
3.1.1 Resultaten	29
3.1.2 De invloed van duur nader onderzocht: matching modellen	36
3.1.3 Conclusies	41
3.1.4 Secundaire uitkomstmaten	43
3.2 Procesevaluatie: vragenlijst	62
3.2.1 Gehanteerde uitkomstmaten	62
3.2.2 Analyse methode	63
3.2.3 Resultaten analyse procesevaluatievragenlijst	63
3.2.4 Conclusies	82
3.3 Focusgroep gesprekken met klantmanagers	83
3.3.1 Onderwerpen	84
3.3.2 Analyse methode	85
3.3.3 Resultaten	85
3.3.4 Samenvatting	96
4 Samenvattende bevindingen en voorlopige conclusies	98
4.1 Opzet van het experiment en werving	99
4.2 Uitvoering van de treatments	100
4.3 Uitstroom naar werk en matching	103
4.4 Niet-werk gerelateerde uitkomstmaten	104
4.5 Tot besluit	105
Literatuur	107
Bijlagen	111

1 Inleiding

Dit eindverslag van het Vertrouwensexperiment Apeldoorn/Epe bespreekt de uitkomsten van het onderzoek naar de effecten van een andere manier van bejegening en ondersteuning van mensen in de bijstand, wat bij aanvang een regelluwe bijstand werd genoemd, vergeleken met de reguliere aanpak. Het onderzoek bestaat uit een product -en een procesevaluatie. De productevaluatie onderzoekt de uitkomsten ofwel effecten van de andere aanpak vergeleken met de standaard aanpak niet alleen op uitstroom naar werk maar ook op welbevinden, gezondheid, keuzevrijheid, sociale participatie en zelfredzaamheid; de procesevaluatie onderzoekt de uitvoering van het experiment bij de gemeente Apeldoorn/Epe. Met andere woorden, hoe verloopt het proces van begeleiding en ondersteuning en wat gebeurt er in de spreekkamers waarin klantmanagers/trajectregisseurs en bijstandsgerechtigden het experiment vorm geven. Daarbij gaat het onder andere om de vraag wat de andere manier van bejegening in vergelijking met de standaardaanpak doet met de klantmanagers en de deelnemers met het oog op de ondersteuningstaken.

Het verslag vormt het vijfde verslag voor de gemeente - eerder is in Februari 2018 een rapport uitgebracht over het profiel van de bijstandsgerechtigden in Apeldoorn/Epe en zijn drie tussenrapportages uitgebracht in maart 2018, september 2018 en maart 2019 – maar het is wel het eerste verslag waarin de uitkomsten van de twee treatment -of interventiegroepen eigen regie en intensieve begeleiding in vergelijking met de standaard treatment worden gerapporteerd. In de tussentijdse verslagen zijn geen effecten van de afzonderlijke treatments gerapporteerd aangezien dat het experiment zou kunnen beïnvloeden waardoor een zuivere meting van de effecten niet meer mogelijk is. Wel zijn in de tussentijdse verslagen, en met name het tweede tussentijdse verslag, de stand van zaken, voortgang en de wijze waarop het uitvoeringsproces van het experiment verloopt weergegeven. Hierdoor is inzicht ontstaan in de ontwikkeling van het proces van experimenteren met de andere aanpak.

Historie

Apeldoorn en Epe zijn op 1 oktober 2017 van start gegaan met wat oorspronkelijk een regelluwe bijstand werd genoemd (Delfos & Oude Alink, 2016) en daarna het Vertrouwensexperiment Apeldoorn/Epe (Muffels & van der Klink, 2017). Apeldoorn en Epe hebben besloten te experimenteren buiten de AMvB¹ om, dus binnen de kaders van de Participatiewet. De gemeente wilde een nieuwe meer effectieve aanpak waarin wederzijds vertrouwen tussen cliënt en gemeente alsook eigen regie, zelfredzaamheid en maatwerk belangrijke uitgangspunten zijn van de begeleiding. Daarom wordt het experiment ook wel “Zelf aan het stuur en op maat” genoemd. Onderdeel daarvan is ook het motiveren en belonen van eigen initiatief van de cliënt als positieve prikkel. Bij de opzet van het experiment is aangesloten bij de opzet van de experimenten in de zes AMvB gemeenten Tilburg, Groningen, Utrecht, Wageningen, Deventer en Nijmegen. Dat betekent dat gekozen is voor een ‘random control trial’ (RCT) design waarbij deelnemers aselect worden toegewezen aan de ‘treatments’ of behandelingsgroepen (zie onderzoeksvoorstel: Muffels & van der Klink, 2017).

¹ AMvB is een acroniem voor “Algemene Maatregel van Bestuur”, welke nodig was om toepassing van het experimenteerartikel 83 in de Participatiewet door gemeenten mogelijk te maken.

Voortrekkersrol Apeldoorn

Apeldoorn was de eerste gemeente en daarmee voortrekker in Nederland om te starten met een experiment met regelluwe bijstand in de Participatiewet buiten de kaders van de AMvB om. Andere gemeenten zoals Epe en Oss haakten al snel aan bij het experiment in Apeldoorn. Daartoe heeft mede op initiatief van de toenmalige wethouder Johan Kruithof Apeldoorn/Epe een soort consortium van gemeenten opgericht waarbij andere gemeenten zich konden aansluiten. In het begin waren ook Geldrop-Mierlo en Renkum van de partij maar zij haakten om verschillende redenen af. Geldrop-Mierlo was evenwel later, vanaf 1 juni 2018, weer van de partij. Niet-AMvB gemeenten hebben meer vrijheidsgraden bij het ontwerp van het experiment dan de AMvB gemeenten doordat zij niet gebonden zijn aan de nogal als beperkend ervaren voorwaarden van artikel 83. Anderzijds hebben zij niet de mogelijkheid om de vrijlatingsbepalingen te verruimen. Nochtans zijn ook de niet-AMvB gemeenten gebonden aan de kaders van de huidige Participatiewet. De toetsing van de rechtmatigheid van de uitkering blijft bestaan en categorale ontheffing van verplichtingen voor specifieke treatment groepen is niet toegestaan.

Landelijke experiment en LOEP

Het onderzoek staat niet op zichzelf maar is onderdeel van een landelijk experiment in de bijstand waarvoor zes gemeenten zich hebben gekwalificeerd. In totaal nemen meer dan 5000 mensen in de bijstand deel aan dit landelijke experiment. Het is daarmee een van de grootste experimenten ooit, niet alleen in Nederland maar wereldwijd. Het landelijke experiment Participatiewet is ook opgenomen in de RCT registry van de American Economic Association (AEA). De onderzoekers van de zes zogenaamde AMvB-gemeenten in Nederland maar ook de vijf niet-AMvB gemeenten hebben zich verenigd in het Landelijk Overleg Experimenten Participatiewet, LOEP genaamd. In dat overleg heeft tijdens de tweejarige periode voortdurende afstemming plaatsgevonden over de opzet van het onderzoek, de verwerking van de data, de analyses, de rapportages en de uitkomsten. De analyses in dit rapport zijn deels dan ook gebaseerd op een gezamenlijk analyseplan dat door de onderzoekers verenigd in LOEP is opgesteld. Daarnaast zijn aanvullende lokale analyses uitgevoerd onder andere vanwege de specifieke lokale aanpak. In Apeldoorn/Epe is bijvoorbeeld een procesevaluatie uitgevoerd welke niet in deze vorm is uitgevoerd in de AMvB gemeenten met uitzondering van het onderzoek van de experimenten in Tilburg en Wageningen waarvoor Tilburg University ook verantwoordelijk is.

Inhoud eindverslag

Dit eindverslag is als volgt opgezet. In deel 1 wordt de opzet van het experiment en het onderzoek nader toegelicht. Dit houdt in dat de beleidsuitgangspunten en filosofie van het experiment, het theoretisch kader, zowel het wetenschappelijke als het beleidstheoretische kader, de werving van de deelnemers, de keuze voor de treatmentgroepen, de wijze van steekproeftrekking, de gestelde onderzoeksvragen en de manier van dataverzameling wordt toegelicht. Vervolgens worden in hoofdstuk 3 de resultaten besproken van de beschrijvende analyses van de product- en procesevaluatie op basis van de vragenlijsten. Hoofdstuk 4 bespreekt summier de belangrijkste voorlopige conclusies van het onderzoek.

Gebruik van CBS microdata en planning rapportage

Dit eindverslag rapporteert over de effecten van het experiment op basis van de informatie die op lokaal niveau bij de gemeenten is verzameld (vragenlijsten deelnemers en procesevaluatie vragenlijsten klantmanagers) dan wel beschikbaar is middels de administratie van de gemeente Apeldoorn/Epe (BUS ofwel bijstandsuitkeringsstatistiek). In overleg met de projectleider(s) van de gemeente Apeldoorn/Epe is afgesproken dat een aanvullende rapportage volgt nadat de effecten op volledige en gedeeltelijke uitstroom naar betaald werk zijn geanalyseerd gebruik makend van de microdata van het CBS. Deze analyse is mede vanwege de Coronacrisis vertraagd en zal nu in juni 2020 klaar zijn. De rapportages van de gemeentelijke rapporten in de AMvB gemeenten vindt plaats op 1 mei 2020 waarna de definitieve rapportage volgt op 1 juni 2020. De zes AMvB-gemeenten maken gebruik van de microdatabestanden van het CBS om de effecten op de zogenaamde primaire uitkomstmaat te analyseren. Deze analyse gebeurt in overleg met het Centraal Planbureau dat door het Ministerie van Sociale Zaken en Werkgelegenheid is verzocht om de evaluatie van deze uitkomstmaat voor haar rekening te nemen. Door gebruik te maken van de door het CBS bewerkte microdata en een iets andere analysemethode kunnen de uitkomsten van de beide analyses verschillen. De onderzoekers zullen deze verschillen in kaart brengen en van toelichting voorzien opdat eenduidige beleidsconclusies mogelijk zijn. In het technische deel van de rapportage dat later volgt wordt een meer gedetailleerde en technische toelichting te geven op de manier waarop het onderzoek is uitgevoerd. Dit gedeelte is bedoeld voor lezers die meer achtergrondinformatie willen hebben van het onderzoek. Dit deel zal onderdeel uitmaken van het definitieve verslag op 1 juni 2020.

1.1 Aanleiding, achtergrond en opzet onderzoek

Wetgeving: de invoering van de Participatiewet per 1 januari 2015

Sinds 1 januari 2015 vallen bijstandsuitkeringen onder de regelingen van de Participatiewet (PW), die de oude Wet werk en bijstand (Wwb) heeft vervangen. Doel van de Participatiewet is om zoveel mogelijk mensen naar betaald werk te helpen of, als dat niet kan, op een andere manier te laten participeren. De wet regelt de uitgifte van bijstandsuitkeringen en geeft een kader voor de regels en verplichtingen die bijstandsgerechtigden moeten naleven. Hoewel de wet de nationale kaders vastlegt, ligt de verantwoordelijkheid voor de uitvoering van de wet bij de gemeenten. Het doel van de wet was om meer mensen sneller te laten uitstromen naar betaald werk (zie SCP rapport, 2019). Na invoering van de wet op 1 januari 2015 is bij meerdere gemeenten de behoefte ontstaan om de doeltreffendheid van de nieuwe wet te onderzoeken, met andere woorden of de Participatiewet de gestelde doelen bereikt; werkt het goed en zo niet, wat en hoe kan het dan beter werken. Doeltreffendheid richtte zich daarbij zowel op de toeleiding naar betaald werk als ook op de bevordering van de participatie in de samenleving. Bij het tweede doel kan gedacht worden aan vrijwilligerswerk, psycho-sociale of andere hulpverlening zodat mensen beter om kunnen gaan met hun persoonlijke en/of sociale problemen en daardoor beter kunnen participeren. In algemene zin vroegen de gemeenten zich af of de P-wet voldoende is toegerust om ondersteuning te kunnen bieden voor verbetering van de sociale participatie door middel van het bevorderen van de zelfredzaamheid, gezondheid en welbevinden van de bijstandsgerechtigde welke hiervoor bepalende factoren zijn. Bovendien vonden een aantal gemeenten de regelingen en verplichtingen in de Participatiewet niet altijd goed aansluiten bij de huidige praktijk. Het stelsel van regels

en verplichtingen in de Participatiewet werd door gemeenten als complex en bureaucratisch ervaren, met weinig ruimte voor maatwerk en vormen van begeleiding waarbij de bijstandsgerechtigde zélf meer regie krijgt. Uiteindelijk zijn 11 gemeenten aan de slag gegaan om te experimenteren met een andere manier van begeleiding waaronder zes gemeenten die na een aanvraagprocedure zijn toegelaten als experiment in het kader van het experimenteer artikel 83 van de Participatiewet. Vijf, waaronder Apeldoorn en Epe, dus buiten de AMvB om.

Wat houdt het onderzoek in?

Het onderzoek bestaat zoals eerder opgemerkt uit twee onderdelen: de product –en procesevaluatie (zie Muffels & van der Klink 2017). De *productevaluatie* is erop gericht na te gaan wat beter werkt voor de cliënt; de nieuwe regelluwe aanpak of de reguliere aanpak die gekenmerkt wordt door een strenge toetsing op de bestaande verplichtingen en de toepassing van maatregelen en kortingen wanneer onverhoopt de bijstandsgerechtigde hieraan niet voldoet. Over een periode van twee jaar zijn deelnemers aan het experiment gevolgd. In het onderzoek wordt er niet alleen gekeken naar of mensen eerder betaald werk vinden maar ook wat het met mensen doet, met hun gezondheid, hun welbevinden, hun zelfvertrouwen en hun zelfredzaamheid. Daartoe vullen deelnemers aan het begin van het experiment, na een jaar en na twee jaar een online vragenlijst in. De eerste meting noemen we de nulmeting en de twee anderen, de eerste en tweede vervolgmeting.

De *procesevaluatie* is erop gericht na te gaan hoe de projectleider(s) en de klantmanagers als professional het proces van implementatie van het experiment in de gemeente percipiëren en waarderen, en welke meer -of minderwaarde de klantmanagers ervaren van de andere bejegening van de cliënt voor de uitoefening van hun taken in de nieuwe aanpak. Daartoe zijn om de acht maanden, in totaal vier, focusgroep gesprekken gehouden met de klantmanagers en de projectleider(s). Ten tweede vult de coach na elke 6-9 maanden op basis van de gesprekken in deze periode voor elk van hen een korte procesevaluatievragenlijst in over de afspraken en bevindingen van de afgelopen periode. In totaal zijn door de coaches drie vragenlijsten per deelnemer ingevuld. De eerste door hen ingevulde procesevaluatievragenlijst vormt tezamen met het eerste focusgroep gesprek met de klantmanagers en projectleider(s) de nulmeting van de procesevaluatie. De vier focusgroep gesprekken daarna en de drie ingevulde vragenlijsten zijn onderdeel van de vervolgmetingen van de procesevaluatie. Vanaf 1 oktober 2018 zijn de klantmanagers vervolgens gestart met de invulling van de procesevaluatie vragenlijsten voor elke klant. De focusgroep gesprekken met de klantmanagers en de projectleiders van Apeldoorn/Epe hebben vervolgens op 18 januari 2018, 27 september 2018, 20 mei 2019 en 6 november 2019 plaatsgevonden.

1.2 Theoretisch kader

Tegen de achtergrond van de vragen die de gemeenten zich stelden over de doeltreffendheid van de nieuwe Participatiewet die per 1 januari 2015 werd ingevoerd en de wens in de gemeente Apeldoorn/Epe om te onderzoeken of een andere aanpak ‘beter werkt’ zijn door de onderzoekers drie treatments uitgewerkt welke zijn gebaseerd op verschillende theoretische inzichten uit de literatuur (zie Muffels & van der Klink, 2017). Twee daarvan, namelijk eigen regie en intensieve begeleiding zijn ook in de opzet van het experiment in Apeldoorn/Epe opgenomen. Voor de derde treatment van ‘extra vrijlating van

verdiensten' bestond er geen wettelijke basis, daarvoor moest gebruik worden gemaakt van artikel 83 welke alleen door de AMvB gemeenten gebruikt mag worden. De theoretische begrippen die aan de basis van de ontwikkelde treatments staan zijn: wederkerigheid, vertrouwen, intrinsieke motivatie, keuzevrijheid, belonen (ook niet-financieel door positieve bejegening), zelfredzaamheid en individueel maatwerk (afgestemd op de wensen en mogelijkheden van de bijstandsgerechtigde). Deze begrippen zijn afgeleid uit de gedragseconomie (belonen, vertrouwen, intrinsieke motivatie), psychologische cognitie - en motivatietheorieën (wederkerigheid, zelfredzaamheid en intrinsieke motivatie) en de capability theorie van Amartya Sen (keuzevrijheid en maatwerk) (zie ook Groot, Muffels en Verlaet, 2019). Deze zullen hieronder nader worden toegelicht. Eerst gaan we in op de achterliggende beleidstheorie die aan de basis heeft gestaan van de vragen die de gemeente zichzelf stelde bij aanvang van het experiment.

Beleidstheorie

In artikel 83 van de Participatiewet wordt als 'primaire uitkomstmaat' van het experiment de uitstroom naar voltijd betaald werk genoemd. Dat betekent dat wanneer mensen minimaal 27 uur gaan werken tegen het minimumloon zij volledig onafhankelijk worden van de uitkering en dus uitstromen uit de bijstand die immers voor een alleenstaande 70% van het minimumloon bedraagt. Het primaire doel van de Participatiewet is zoals opgemerkt om meer mensen sneller naar voltijds betaald werk toe te leiden en daardoor participatie in de samenleving te bevorderen. Lukt dit niet dan zijn tussenstapjes op weg naar dit doel (activering) evenwel ook mogelijk (deeltijdwerk, vrijwilligerswerk, cursussen, training en opleiding, beschut werk, proefplaatsingen, aanpassingen werkplek en loonkostensubsidie voor(gedeeltelijk) arbeidsongeschikten). Loonkostensubsidie en beschut werk zijn nieuw in de P-wet. In de praktijk bleek dit doel zelfs in de herstelperiode van de arbeidsmarkt vanaf 2015 dus na de economische crisis moeilijk haalbaar. De arbeidsmarkt herstelde zich langzaam in de eerste jaren na de ernstigste naoorlogse economische recessie die duurde van 2008 tot einde 2014. Vanaf 2017 ging het beter en veranderde de arbeidsmarkt hetgeen in de jaren 2018 en 2019 leidde tot een erg krappe arbeidsmarkt. Het aantal tijdelijke banen bleef hoog en zelfs stijgen maar tegelijkertijd nam het aantal vaste banen toe. Dit gebeurde in de tijd dat het experiment van start ging. Het experiment vond dus plaats in een gunstig economisch tij waardoor met name jongere bijstandsgerechtigden een plekje konden vinden op de arbeidsmarkt. Voor velen in de bijstand (mensen met slechte gezondheid, ouderen, geen startkwalificatie) was dat ook in deze periode niet mogelijk.

Werkbedrijf Lucrato versus Activering en Inkomen

In Apeldoorn/Epe is in 2017 gestart met een reorganisatie in de uitvoering die per 1 januari 2018 in werking is getreden. Het motto van deze reorganisatie was "Iedereen doet mee" en de intentie was om de uitvoering meer 'mensgericht' te maken in plaats van 'systeem' of 'wet' gericht (Beijer, 2017). Dat laatste betekende bijvoorbeeld dat er meer aandacht is voor een vraaggerichte aanpak (wat wil de bijstandsgerechtigde?) en individueel maatwerk. Daardoor veranderde de reguliere dienstverlening en ging deze wat al meer lijken op de treatments in het experiment. Onderdeel van deze reorganisatie was dat mensen bij de intake voor de PW uitkering ingedeeld werden in mensen met een korte dan wel langere afstand tot de arbeidsmarkt. Degenen met de beste kansen op uitstroom naar werk worden sinds de reorganisatie begeleid en ondersteund door de consultants van het Werkbedrijf Lucrato, een externe organisatie. Degenen waarvoor betaald werk op korte termijn, binnen een jaar, geen optie is worden

begeleid door trajectregisseurs bij de reguliere dienstverlening van Activering en Inkomen. Voor de eerste groep, de Lucrato groep, die veelal korter dan een jaar in de uitkering zit is de ondersteuning erop gericht mensen snel naar voltijds werk te begeleiden. Daartoe kan door de Lucrato consultants een beroep worden gedaan op het volledige instrumentarium van Activering en Inkomen, inclusief de schuldhelpverlening. Voor de A&I groep is het beleid erop gericht om via tussenstapjes hen voor te bereiden op betaald werk. In Apeldoorn/Epe werd het beleid voor de start van het experiment einde 2017 gekenmerkt door een 'stevige' aanpak, met een strenge toetsing op rechtmatigheid en voldoen aan de verplichtingen, maar wel met actieve ondersteuning op basis van een breed palet aan trajecten. Tijdens het experiment is het beleid onder invloed van de reorganisatie meer aandacht gaan besteden aan de effectiviteit van de begeleiding en daarmee aan het doelmatigheidsaspect. Daarbij is in de loop der jaren het accent ook meer komen te liggen op praktische ondersteuning zoals bij schuldhelpverlening. Er wordt vanuit het management en beleid actiever ingezet op deze praktische ondersteuning en met name schuldhelpverlening, hetgeen naar eigen zeggen van de trajectregisseurs veel tijd vraagt. Deze veranderingen impliceren een verandering in de beleidstheorie op basis waarvan het beleid is geformuleerd. Daarmee veranderde echter ook de reguliere dienstverlening en lijkt de gewijzigde aanpak meer op de treatments die in het experiment zijn uitgevoerd. We verwachten dat dit effect zal hebben op de uitkomsten van het experiment waarin de treatments worden vergeleken met de reguliere dienstverlening.

Bij Lucrato is het hoofddoel snelle toeleiding naar betaald werk terwijl bij Activering en Inkomen tussenstapjes mogelijk zijn (deeltijdwerk, vrijwilligerswerk, beschut werk, proefplaatsingen, loonkostensubsidie voor (gedeeltelijk) arbeidsongeschikten, opleiding en training, of inschakeling van andere hulpverlening). Wat het eerste doel betreft, toeleiding naar betaald werk, bleek dit doel zelfs in de herstelperiode van de arbeidsmarkt vanaf 2015 dus na de economische crisis, moeilijk haalbaar. De arbeidsmarkt herstelde zich langzaam in de eerste jaren na de ernstigste naoorlogse economische recessie die duurde van 2008 tot einde 2014. Vanaf die tijd maar vooral vanaf 2017 ging het beter en veranderde de arbeidsmarkt hetgeen in de jaren 2018 en 2019 leidde tot een erg krappe arbeidsmarkt. Het aantal tijdelijke banen bleef hoog en zelfs stijgen maar tegelijkertijd nam het aantal vaste banen toe. Dit gebeurde in de tijd dat het experiment van start ging. Het experiment vond dus plaats in een gunstig economisch tij waardoor met name jongere bijstandsgerechtigden een plekje konden vinden op de arbeidsmarkt. Voor velen in de bijstand (mensen met slechte gezondheid, ouderen, geen startkwalificatie) was dat ook in deze periode niet mogelijk. De meerderheid van de mensen in de bijstand betreft immers een zeer kwetsbare groep waarvan slechts een klein deel kan worden toegeleid naar betaald werk (op basis van CBS cijfers berekenden we een percentage van 7,5% per jaar of 15% over 2 jaar in Apeldoorn/Epe; zie Muffels e.a. 2018). Bovendien is de kans om in een tijdelijke baan terecht te komen groot (80%) waardoor de kans op terugkeer in de bijstand op korte termijn (binnen 2 jaar) ook groot is. In Apeldoorn/Epe is deze kans tussen 2017 en 2019 ongeveer 20%. Drie van de vier bijstandsgerechtigde in Apeldoorn/Epe hebben geen startkwalificatie (MBO-niveau 2) en zijn daardoor voor de meeste werkgevers onvoldoende geschikt voor een baan. Daarnaast vinden velen een tijdelijke baan maar met een betrekkelijk gering aantal arbeidsuren per week waardoor ze deels afhankelijk blijven van de bijstand. Voor zover mensen gedeeltelijk of volledig uitstromen naar betaald werk lukt dat alleen met (intensieve) persoonlijke begeleiding en maatwerk (zie o.a. Muffels e.a. 2016).

De stap naar betaald werk is voor een groot deel van de bijstandsgerechtigden dus te groot; voor velen zijn tussenstapjes nodig voordat een terugkeer in betaald werk mogelijk is, voor anderen in het zogenaamde “granieten bestand” is uitstroom naar betaald werk een onhaalbare kaart zelfs op de wat langere termijn. Zij hebben speciale ondersteuning nodig met een specifieke maatgerichte aanpak in verband met hun persoonlijke en maatschappelijke problemen en problemen met hun fysieke of mentale gezondheid. Op basis van de informatie in de bijstandsuitkeringsstatistiek (BUS) bleek dat ruim tweederde of 67% van de mensen in de bijstand in Apeldoorn/Epe met gezondheidsproblemen kampten waarvan een derde met psychische problemen (zie eerste tussenrapportage: Muffels e.a. 2018).

In de eerste tussenrapportage in maart 2018 is met behulp van latente klassenanalyse² een analyse gemaakt van de deelnemers aan het experiment op basis van informatie in de eerste vragenlijst. Het zoek- en sollicitatiegedrag van de deelnemer is geanalyseerd plus de belemmeringen die zij ervaren onder andere vanwege hun gezondheidsproblemen en een laag niveau van arbeidsgeschiktheid (inclusief hun kwalificaties). Later is deze analyse herhaald op alle mensen die de nulmeting hebben ingevuld en waarvoor de informatie beschikbaar was (n=440). Het beeld bleef hetzelfde. Zes kenmerken van de deelnemers in Apeldoorn/Epe zijn gebruikt om de profielen te maken. Deze hebben allen te maken met het zoeken naar werk en de perceptie van de mate van arbeidsgeschiktheid inclusief de fysieke en mentale beperkingen die mensen ervaren om te gaan werken. De beste oplossing bij de analyse was een indeling in vier groepen. De eerste groep noemen we de ‘actief op zoek, arbeidsgeschikte’ groep en bestaat uit mensen die aan het werk willen en actief op zoek zijn. Dat is het geval voor 39% van de deelnemers (n=173). Zij zijn allen actief op zoek naar betaald werk (100%) en zien daar ook nog mogelijkheden voor. Dan is er een groep die niet meer actief zoekt en berustend of ontmoedigd is vanwege het zoeken naar werk en het niet slagen daarvan (n=62 of 14%). Tenslotte zijn er twee groepen die tezamen bijna de helft van de populatie uitmaken (46%); zij zoeken niet actief meer en zien vanwege de ernst van hun gezondheidsproblemen nog weinig kansen voor het vinden van betaald werk. De eerste groep is passief vanwege fysieke gezondheidsproblemen (14%) en de tweede vanwege mentale gezondheidsproblemen (32%). Deze tweede groep die zich vanwege mentale gezondheidsproblemen arbeidsongeschikt acht is relatief groot in Apeldoorn/Epe. Deze groep vraagt om een specifieke aanpak en een passende beleidstheorie of filosofie om deze adequaat te begeleiden naar participatiemogelijkheden inclusief (aangepast) werk. De filosofie achter het vertrouwensexperiment alsook het design van de treatments die eruit voortkwam was niet specifiek ingericht op deze problematische groep. Wel werd in het onderzoeksvoorstel gesteld dat ook deze groep baat kan hebben bij de voorgestelde treatment van intensieve begeleiding omdat deze meer vraaggericht is en gebaseerd is op maatwerk. In het

²Latente klasse analyse probeert aan de hand van een aantal waargenomen variabelen (bevroegd in de nulmeting) een achterliggende) niet waargenomen variabele te vinden (zie <https://nl.wikipedia.org/wiki/Latenteklassemodel>). Minimaal bestaat deze variabele uit twee categorieën verwijzend naar twee latente klassen. Een latente klasse wordt in de analyse weergegeven door voorwaardelijke kansen die voor iedere mogelijke score op de waargenomen variabelen wordt berekend. Deze kans geeft aan hoe waarschijnlijk een deelnemer die tot een van de klassen behoort een bepaalde score heeft op de waargenomen variabelen. De oplossing is technisch gezien uitstekend: er is een sterk onderscheid tussen de klassen en het aantal classificatiefouten is gering (6%). De analyses zijn uitgevoerd met het programma Latent Gold 5.1.

onderzoeksvoorstel (Muffels & Klink, 2017) is daarom verwezen naar experimentele studies van IPS (“individual placement support”)-projecten en ES (“employment support”)-programma’s die met name voor deze groep soelaas kunnen bieden (Bond et al. 2010; Marshall et al. 2014). Dergelijke IPS/ES-programma’s richten zich evenwel niet alleen op de aanbodzijde maar ook op de vraagzijde van de arbeidsmarkt hetgeen beter past bij een werkgeversaanpak waarin bijvoorbeeld ook mogelijkheden voor aanpassing van het werk worden meegenomen. Het vertrouwensexperiment beperkte zich evenwel tot de aanbodzijde.

Vanuit het besef van een problematische arbeidsmarkt voor deze kwetsbare groepen hebben de gemeenten bij aanvang van het experiment de wens geuit om naast de primaire uitkomstmaat van volledige uitstroom naar betaald werk ook een aantal secundaire uitkomstmaten te gebruiken om te bepalen of het experiment al dan niet werkt voor de deelnemers. Juist voor mensen waarvoor de stap naar betaald werk te groot is kunnen deze secundaire uitkomsten van groot belang zijn om later ook andere stappen te kunnen zetten. De uitkomstmaten die de gemeente Apeldoorn/Epe heeft gekozen zijn: vertrouwen, keuzevrijheid, bevordering van zelfredzaamheid, welbevinden en gezondheid (fysiek en mentaal) en sociale participatie. Daarnaast werd belang gehecht aan een betere relatie tussen burger en overheid door een andere bejegening van de burger uitgaande van vertrouwen, individueel maatwerk en aandacht in plaats van ‘loslaten’, ‘weinig of geen contact’ en wantrouwen welke begrippen in de ogen van de gemeente beter passen bij de reguliere aanpak. De ‘mensgerichte’ aanpak die vanaf 1 januari 2018 in gang is gezet heeft evenwel de reguliere dienstverlening veranderd waardoor deze meer is gaan lijken op de experimentele dienstverlening.

Veel gemeenten die vanaf de invoering van de nieuwe P-wet werden geconfronteerd met de volledige verantwoordelijkheid voor de financiering van de bijstandsuitkeringen middels de BUIG-middelen zochten een weg om de wet doeltreffender uit te voeren. Een stringente toetsing op verplichtingen en voorwaarden was een van de manieren om dat te doen. Sommige gemeenten zijn die weg op gegaan, anderen zochten juist naar mogelijkheden om de effectiviteit van de re-integratie te verbeteren middels een minder strikte toepassing van de arbeids- en re-integratieplicht en meer ruimte voor eigen regie van de klant. Die weg is Apeldoorn/Epe opgegaan in de vorm van het inrichten van het vertrouwensexperiment waarin naast eigen regie en het minder stringent toepassen van de verplichtingen een intensievere en meer op maat gesneden begeleiding en ondersteuning wordt aangeboden vooral ook aan mensen in het zogenaamde “zittende” bestand. Daarom wordt het experiment ook wel aangeduid met ‘zelf aan het stuur’ en ‘begeleiding op maat’. Daarnaast heeft ook de reorganisatie per 1 januari 2018 in de uitvoering een rol gespeeld bij de verandering van de reguliere dienstverlening vanuit het perspectief van ‘iedereen doet mee’ en een ‘mensgerichte’ aanpak.

Wetenschappelijk kader

Het experiment is theoretisch ontleend aan een viertal inzichten, die ontleend zijn aan een brede onderzoeksliteratuur in de gedragseconomie, de empirische sociologie en de cognitie -en motivatiepsychologie (zie Groot, Muffels en Verlaet, 2018):

- Het eerste inzicht gaat over recente onderzoeksresultaten over de invloed van armoede op de 'mindset' ofwel geestesgesteldheid van mensen. Onderzoek in dit relatief nieuwe onderzoeksgebied laat zien dat (financiële) schaarste en stress door armoede de cognitieve hulpbronnen van mensen verminderen (Manietal, 2013; Mullainathan en Shafir, 2013). Men heeft alle energie, geestkracht en inventiviteit nodig om de eindjes voor henzelf, eventuele partner en kinderen aan elkaar te knopen door het ontwikkelen van zogenaamde huishoudbezuinigingsstrategieën om in de eigen behoeften en die van het huishouden te kunnen voorzien. Indien financiële schaarste en het voldoen aan bijstandsverplichtingen een groot deel van de cognitieve hulpbronnen van mensen innemen, is er weinig ruimte voor belangrijke en cognitief uitdagende taken, zoals het bijscholen voor een andere baan, het onderhouden van het sociaal netwerk of het actief zoeken naar betaald werk (Groot, Muffels en Verlaet, 2019).
- Het tweede inzicht is afkomstig uit de gedragseconomie en gaat over de invloed van instituties op de arbeidsmarkt en in de sociale zekerheid alsmede de impliciete waarden die zij vertegenwoordigen op het gedrag van mensen. De impliciete waarden die ten grondslag liggen aan deze instituties en die tot uiting komen in uitkeringsvoorwaarden, regels en verplichtingen beïnvloeden het gedrag van de gebruikers in de vorm van bijvoorbeeld het zoekgedrag op de arbeidsmarkt, maar ook de keuze tussen werken en vrije tijd in de vorm van bijvoorbeeld het aantal uren dat men wil werken. Daarnaast beïnvloeden zij andere economische beslissingen in de sfeer van consumptie en bestedingen, huisvesting, onderwijs en gezondheid niet alleen van henzelf maar ook van anderen in het huishouden waaronder ook eventuele kinderen. Onderliggende waarden zijn bijvoorbeeld reciprociteit (wederkerigheid, 'tit for tat') en vertrouwen. Reciprociteit betekent dat individuen een goede behandeling of het ontvangen van vertrouwen (een investering in sociale relaties), belonen door bijvoorbeeld zich extra in te spannen (positieve reciprociteit), terwijl ze het omgekeerde doen als ze slecht worden behandeld of op basis van wantrouwen worden bejegend (negatieve reciprociteit) (Fehr en Schmidt, 2003). Negatieve prikkels, zoals opgelegde financiële sancties en kortingen zijn niet noodzakelijkerwijze de beste manier om mensen te prikkelen mee te werken en zich aan de regels te houden. Economisch en sociologisch onderzoek laat zien dat er weliswaar bescheiden korte termijn effecten optreden (men accepteert sneller werk ook al is de 'fit' minder) maar dat de effecten op de langere termijn onzeker zijn bijvoorbeeld doordat een slechte 'fit' tussen mens en werk op de iets langere termijn welbevinden, gezondheid en de loopbaan negatief beïnvloeden (zie SCP, 2019; Blonk, 2018). Bevindingen uit de experimentele economie laten ook zien dat mensen in ruil voor het vertrouwen dat zij ontvangen extra gemotiveerd zijn en meer hun best doen voor hun taak en daarmee dus ook degene die hen vertrouwt beloont (Groot, Muffels en Verlaet, 2019). Op deze manier leidt vertrouwen dus tot gevoelens van positieve reciprociteit en daardoor ook tot blijvende inzet en toegenomen productiviteit (Bohnet et al., 2001).
- Het derde inzicht is afkomstig uit de psychologische motivatie theorie en leert ons dat extrinsieke prikkels de intrinsieke motivatie kunnen wegnemen (Frey en Jegen, 2001). De zelfbeschikkingstheorie (Deci en Ryan, 1985) stelt dat intrinsiek gemotiveerde mensen een activiteit ondernemen omdat ze het plezierig en interessant vinden, waardoor ze meer effectiviteit en volharding laten zien in hun gedrag en een verbeterd welbevinden (Ryan et al.,

1997). Deze theorie stelt ook dat het geven van vertrouwen aan mensen een gevoel van zelfsturing oplevert, die weer effect heeft op het werkzoekgedrag en duurzame werkgelegenheid. Voorgaand onderzoek wijst ook uit dat de intrinsieke motivatie kan worden versterkt door een activiteit als keuze aan te bieden in plaats van als controlemiddel (Groot, Muffels en Verlaet, 2019).

- Het laatste inzicht gaat over 'vrije keuze' en komt voort uit Sen's 'capability theorie'. Binnen deze theorie zijn 'capabilities' de keuze opties die mensen worden aangeboden om de dingen te doen die zij belangrijk vinden in hun eigen leven. Op deze manier krijgen mensen kansen aangeboden waardoor ze persoonlijke autonomie en zelfvertrouwen ontwikkelen en daarmee hun welbevinden kunnen vergroten (Sen, 2004; 2009). Op basis van de inzichten van Sen wordt in recente theorievorming rond duurzame inzetbaarheid het waarde-aspect van werk benadrukt alsook het belang van eigen regie (Van der Klink et al., 2016). De focus op keuzemogelijkheden en eigen regie vergroot dus ook de intrinsieke motivatie om te gaan werken.

Deze vier inzichten vormen de uitgangspunten voor de manier waarop de eerder genoemde alternatieve interventie-benaderingen (of treatments) zijn vormgegeven binnen het vertrouwensexperiment. Tevens zijn ze inspiratiebron geweest voor het opstellen van vragen in de vragenlijsten aan de deelnemers en de coaches. De inzichten zijn niet alleen relevant voor de productevaluatie waarin het effect wordt onderzocht op arbeidsinschakeling en op welbevinden, gezondheid en sociale participatie maar ook voor de procesevaluatie omdat de procesevaluatie onderzoekt in hoeverre sturing op zelfredzaamheid, intrinsieke motivatie en keuzevrijheid ook beter werkt in de uitvoering. We hebben aan het begin de volgende hypothesen geformuleerd.

Hypothesen

- Op grond van gedragseconomische inzichten verwachten we dat positieve prikkels in de vorm van bijvoorbeeld een andere, positievere en meer ondersteunende bejegening met minder verplichtingen en gebaseerd op vertrouwen leidt tot een grotere motivatie en tot meer coöperatief gedrag bij de cliënten hetgeen zich uit in een grotere bereidheid tot deelname alsook inzet om de re-integratie tot een succes te maken. We verwachten positieve effecten op motivatie en daarmee op participatie en arbeidsinschakeling in deze groepen in vergelijking met de controlegroep.
- Omgekeerd verwachten we op grond van het wetenschappelijk empirische onderzoek dat een systeem van strenge handhaving, monitoring en sanctionering zoals bij de controle –en referentiegroep positieve effecten kan hebben op inschakeling in werk maar dat de vraag is of deze effecten korte termijn effecten zijn en dat duurzame uitstroom mogelijk minder zal zijn bij deze groep dan bij de treatment groepen
- Een streng systeem van controle en handhaving kan negatieve effecten hebben op de motivatie(motivatietheorieën) en de mentale belasting en stress die het systeem van controle en straffen met zich meebrengt met negatieve gevolgen voor welbevinden en de (mentale) gezondheid.

- In overeenstemming met de uitkomsten van het Fit4Work programma en de uitkomsten van review studies naar de effecten van IPS (individual placement support) en ES ('employment support') interventies verwachten we sterke effecten van intensieve begeleiding op maat op de arbeidsinschakeling van vooral de meest kwetsbare groepen in de bijstandspopulatie (het zittende bestand).
- De focus op eigen regie en zelfredzaamheid geeft mensen handelings- en keuzevrijheid terwijl intensieve begeleiding en maatwerk hen extra keuzemogelijkheden en kansen biedt. Hierdoor neemt hun zelfvertrouwen en welbevinden toe waardoor het zoekgedrag toeneemt alsook de intrinsieke motivatie om te gaan werken hetgeen uiteindelijk ook de arbeidsinschakeling bevordert. Dit verwijst naar Sen's 'capability' theorie en Ryan en Deci's zelfdeterminatie theorie.

Procesevaluatie

De procesevaluatie beoogt inzicht te krijgen in hoe het proces van experimenteren verloopt in de spreekkamer van de professional waar de ontmoeting met de burger plaatsvindt. Hoe vertaal je de treatments naar de uitvoeringspraktijk; wat betekent eigen regie en hoe intensief is intensief in de treatment intensieve begeleiding? Hoe worden de treatments uitgevoerd, op welke manier wordt een vertrouwensband met de deelnemer opgebouwd, welke condities belemmeren een goede uitvoering van de treatments etc.? De beschikbare wetenschappelijke literatuur over 'wat en waarom werkt' in het re-integratie beleid is beperkt. In de oratie van Blonk (2018) wordt gesteld dat de meeste studies gericht zijn op beleid en beleidsevaluatie meer dan op het begrijpen van de individuele gedragsverandering en het bieden van handvatten aan de professional. Evaluatiestudies van re-integratie beleid laten zien dat de effecten van dat beleid (bemiddeling, werkcreatie, loonkostensubsidies) niet erg groot zijn, en veelal korte termijn effecten laten zien (zie de overzichtsstudies van Card, Kluve en Weber, 2010; 2015). De effecten van het activerend arbeidsmarktbeleid bedragen gemiddeld over alle studies voor de middellange termijn (12 tot 24 maanden) 11,5%. Voor experimenteel onderzoek bedraagt het gemiddeld percentage 5,6%. Dit betreft het gemiddelde van begeleiding bij het zoeken naar werk, werk creatie in publieke sector, controle/sancties, subsidies voor private werkgelegenheid en training. Dit betreft dus zowel mensen in de werkloosheidsverzekering als in de bijstand. De studies tonen wel aan dat training en ontwikkeling dus investeren in menselijk kapitaal effectiever is dan de andere instrumenten voor de wat langere termijn. Een recente studie van een soortgelijk 2-jarig experiment met intensieve begeleiding en maatwerk in Lausanne in Zwitserland tussen 2015-2017 laat zien dat het effect op uitstroom naar werk gemiddeld zo'n 9% bedraagt over de gehele periode (Bigotta, Bonoli, Fati, Lalive, and Oesch 2018). Internationaal onderzoek toont ook aan dat Intensieve begeleiding van mensen met een psychische beperking werkt en effectief om deze mensen naar werk toe te leiden. De effecten hiervan zijn groter dan de hierboven genoemde effecten en variëren gemiddeld tussen 10 tot 25% (gemiddeld 15%). De meest sprekende voorbeelden van succesvolle begeleiding naar werk van mensen met (mentale en fysieke) gezondheidsbeperkingen komen dan ook uit de psychologische literatuur rond "Individual Placement Support" en "Employment Support" (zie Bond et al. 2010; Marshall et al. 2014). Daarin wordt specifiek en gericht individueel maatwerk geboden (zowel aan de aanbod als de vraag -of werkgeverskant) welke leidt tot de genoemde additionele uitstroompercentages.

Blonk (2018) stelt in zijn oratie daarom het volgende: “Je moet dus investeren in mensen en dat vooral volhouden. Dat is het beste dat je kunt doen. Maar we weten ook dat niet elke training helpt en ook niet voor iedereen in gelijke mate. Om daar meer grip op te krijgen is inzicht nodig in de werkzame bestanddelen van en de condities rond een training. Niet alleen wat werkt, maar waarom, zodat de professional onderbouwde keuzes kan maken wat wordt ingezet. Die kennis is nog maar beperkt beschikbaar. Er is wat dat betreft nog steeds een hoog ‘black box’ gehalte. Om dat op te lossen zijn meer studies nodig waarin naar gedragsaspecten wordt gekeken. Aspecten zoals kwaliteit van zoekgedrag, naar de intensiteit daarvan en de focus, naar vertrouwen in eigen in dat zoekgedrag en naar vaardigheden die zelfregie ondersteunen en hoe op deze aspecten te interveniëren. Wat dat betreft zijn we nog maar net begonnen.” (citaat oratie Blonk, 2018). Vervolgens verwijst Blonk naar studies van Liu, Wang en Huang uit 2014 waarin wordt onderzocht wat werkt bij een aantal specifieke sollicitatietrainingen die hebben geleid tot een driemaal grotere kans om werk te vinden voor de deelnemers. Het is volgens Blonk de combinatie van motivatie en gedrag die bepalend zijn voor het succes van deze trainingen. Deze overwegingen hebben aan de basis gestaan van de procesevaluatie en de invulling van de ‘skills’ training die aan klantmanagers is gegeven om deelnemers te leren zelfredzaam te zijn of te worden. De basis van de training wordt gevormd door de theorie van ‘planned behavior’ ofwel het integratieve gedragsmodel van Fishbein & Ajzen (2010). In het model (Fig. 1.1) staat zoekgedrag naar werk centraal. Het gaat om intenties ofwel motivatie en vaardigheden. De motivatie is afhankelijk van het belang van het te bereiken doel, de ervaren druk van de omgeving en hoe je daarmee omgaat en de verwachting dat je het doel kunt halen. Bij vaardigheden gaat het om vaardigheden die je ten dienste staan bij het eigen regie nemen en houden over je leven, over zelfredzaamheid, volhouden en niet bij de pakken te gaan neerzitten ondanks je belemmeringen. Deze vaardigheden zijn te leren.



Fig. 1.1. Het integratieve gedragsmodel van Fishbein & Ajzen (2010).

Bron: Oratie Blonk (2018):

De laatste component in het model zijn de mentale en fysieke beperkingen maar ook kenmerken zoals leeftijd, nationaliteit of culturele achtergrond. Het omgaan met deze beperkingen en kenmerken is onderdeel van zelfredzaamheid en eigen regie. De training is een aantal keer in Apeldoorn/Epe gegeven. In de focusgroep gesprekken met de klantmanagers is de betekenis en rol van de training in het uitvoeringsproces voortdurend besproken. In de verslaglegging van deze gesprekken wordt hier nader op ingegaan. Vanuit wetenschappelijk oogpunt is nu vooral van belang om te toetsen of als de uitvoering niet tot het gewenste resultaat leidt dit het gevolg is een 'theory failure', de uitgangspunten zijn niet goed geformuleerd of een 'uitvoering failure', de treatments zijn niet conform de onderliggende uitgangspunten uitgevoerd. De vragenlijsten bij de coaches en de focusgroep gesprekken hebben tot doel om hier meer inzicht in te krijgen. Daarbij analyseren we in samenhang de invloed van contextuele factoren (politiek, gemeentelijke uitvoering), het gedrag van de coaches en deelnemers, de bedoelde en onbedoelde effecten van de treatments, en de uitkomsten en effecten.

1.3 Onderzoeksvragen en hypothesen

Het doel van het onderzoek is om te achterhalen wat het effect is van de twee alternatieve interventies of aanpakken op de gekozen uitkomstmaten in vergelijking met de standaard aanpak. Het idee is dat door een andere meer positieve bejegening (inhoudende eerder belonen dan straffen) en een andere manier van het 'gesprek' aangaan, meer vanuit intrinsieke motivatie en vertrouwen dan vanuit extrinsieke prikkels en wantrouwen, het gedrag positief wordt beïnvloed in de richting van meer eigen initiatief en zelfredzaamheid. Uiteindelijk wordt daarmee ook het hoofddoel van de P-wet om meer mensen sneller toe te leiden naar betaalde arbeid beter bereikt.

Wetenschappelijk gezien zijn de experimenten interessant omdat er nog weinig bekend is over de effectiviteit van een alternatief beleid van motiveren en belonen (positieve prikkels) in plaats van een beleid van straffen en sanctioneren (negatieve prikkels). De doeltreffendheid van het beleid is voor dit onderzoek breed gedefinieerd. Dat houdt in dat er wordt gekeken naar effecten op vijf domeinen: (1) arbeidsparticipatie, (2) sociale participatie en vertrouwen, (3) gezondheid en welbevinden en (5) de financiële situatie van de deelnemers. De doeltreffendheid wordt vastgesteld door de twee alternatieve interventies te vergelijken met de standaardaanpak. Hieruit resulteren een aantal beleids- en onderzoeksvragen:

Beleid -en onderzoeksvragen

- In het experiment worden de verplichtingen, die gekoppeld zijn aan de inspanningen die van cliënten worden verwacht om te participeren en re-integreren, bijvoorbeeld de sollicitatie -en re-integratieplicht versoepeld (met name ook in de eigen regiegroep). De vraag voor beleid is dan wat is het effect van deze ontheffingen op de gekozen uitkomstmaten?
- Een tweede beleidsvraag betreft het effect op de arbeidsinschakeling wanneer mensen op een positieve vraaggerichte manier worden ondersteund en intensiever en op maat worden begeleid (met name intensieve begeleidingsgroep).

- Een derde beleidsvraag voor de gemeente was of mensen die deelnemen aan het experiment meer ruimte ontwikkelen om eigen initiatief te ontplooiën, meer in brede zin gaan participeren, en wat het effect hiervan is op de arbeidsinschakeling (eigen regie, intensieve begeleiding).
- Een belangrijke vraag voor de gemeente was wat de neveneffecten zijn van de ruimere handelingsvrijheid op maatschappelijke participatie, maar ook op zelfredzaamheid, welbevinden en gezondheid (eigen regie, intensieve begeleiding).
- In het experiment wordt onderzocht wat de uitkomsten over de tijd gezien zijn van meer handelingsvrijheid en een intensievere begeleiding op maat. Dat vraagt een andere manier van ondersteuning en andere kwalificaties van de klantmanagers. De klantmanagers krijgen ook meer tijd (lagere caseload) om de deelnemer te begeleiden. De gemeente wil graag weten hoe dit proces verloopt. Worden de treatments uitgevoerd zoals beoogd en in welke mate zijn ze onderscheidend? Hoe verloopt de interactie met de klant? Wat werkt in de uitvoering, in de spreekkamers beter en wat werkt niet ook vanuit het perspectief van de klant? Welke condities, aspecten maken het verschil en wat betekent dit voor de implementatie van de uitkomsten (procesevaluatie).

Analysemodel productevaluatie

Het integratieve (zoek)gedragsmodel van Fishbein en Ajzèn fungeerde als onderliggende theoretische basis voor de wijze van uitvoering van de treatments. Daarbij staat sturing op intrinsieke motivatie en gedrag centraal. In de productevaluatie wordt vervolgens onderzocht welke effecten de treatments op de gekozen uitkomstmaten hebben, zoals voltijdse of deeltijdse uitstroom naar werk, welbevinden, gezondheid, sociale participatie en zelfredzaamheid ervan uitgaande dat de treatments conform de theoretische uitgangspunten zijn uitgevoerd. Daarbij wordt rekening gehouden met factoren die deze effecten beïnvloeden zoals de sociaaleconomische en culturele context, de positie op de arbeidsmarkt, de beleving en keuzemogelijkheden en het zoekgedrag. Dit wordt in Figuur 1.2 weergegeven.

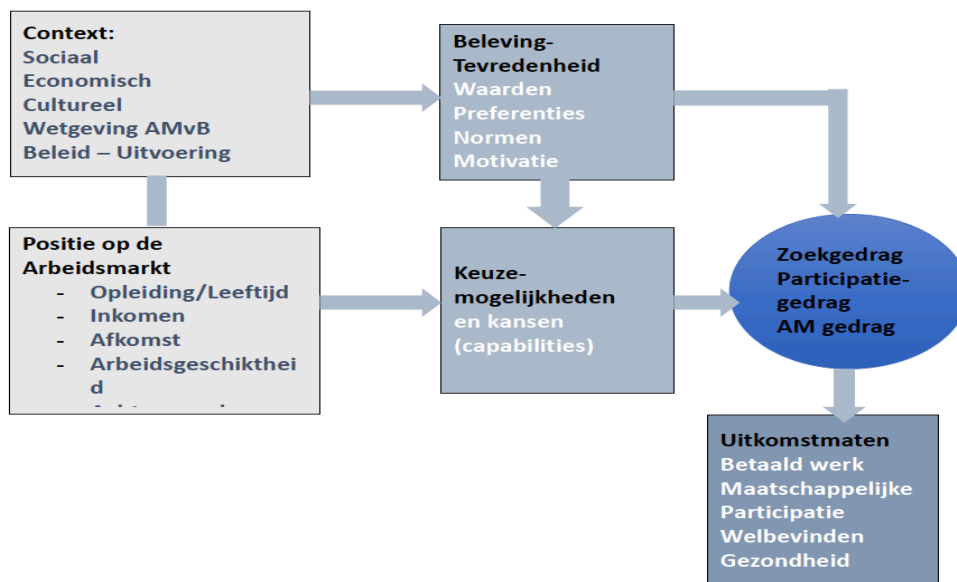


Fig. 1.2. Analysemodel onderzoek effecten Vertrouwensexperimenten PW

2 Opzet van het onderzoek

2.1 Experimentgroepen

Apeldoorn/Epe kent evenals de andere niet-AMvB gemeenten waarvoor Tilburg University de evaluatie verricht drie treatment of interventiegroepen: 1. Eigen regie; 2. Maatwerk/intensieve begeleiding en 3. de reguliere ofwel standaard aanpak (de controlegroep).

1. *Eigen regie of 'zelf aan het stuur'*. Dit is de groep waarbij de focus ligt op het minder stringent toepassen van de verplichtingen en mensen wordt geleerd zelfredzaam te zijn en zelf de regie te nemen om de problemen aan te pakken. Van deze groep wordt verwacht dat zij zelfredzaam zijn of worden om terugkeer naar werk te realiseren, door de deelnemers te leren hoe ze proactief en met zelfvertrouwen (en dus zelfsturend) moeten handelen. Bij deze groep zijn de verplichtingen van de sollicitatie en re-integratieplicht minder strikt toegepast, om zo meer ruimte te laten voor eigen regie van de klant. Sancties worden in deze groep slechts in uitzonderlijke gevallen toegepast. Voor het overige blijft alles zoals in het normale regime, bijvoorbeeld bijverdiensten worden hetzelfde behandeld als voor bijstandsgerechtigden die niet deelnemen aan het experiment.
2. *Op maat*: dit is de groep die individueel maatwerk ontvangt en extra persoonlijke begeleiding krijgt. Bij de intensiveringsgroep blijven alle regels en regelgeving van het bestaande bijstandsregime gelden, waaronder de toepassing van de re-integratieplicht. Wel ontvangt deze groep extra hulp en/of meer contact door middel van op maat gemaakte begeleiding en intensieve bemiddeling van de gemeente. Gemiddeld hebben deze deelnemers twee keer zoveel contactmomenten met de gemeente dan de deelnemers in de controlegroep.
3. *Regulier*: dit is de groep die de standaard aanpak krijgt. Voor hen gelden de bestaande strenge aanvraag- en re-integratieverplichtingen.

'Hawthorne' effecten

De reguliere/standaardaanpak is vanwege de in 2017 in gang gezette reorganisatie van de uitvoering in de loop van de tijd veranderd zoals we eerder constateerden. Dit heeft te maken met autonome beleidsontwikkelingen maar wellicht ook met uitstralingseffecten van het experiment op de deelnemers en vooral ook begeleiders in de reguliere aanpak omdat de consultants/trajectregisseurs onderdeel uitmaken van het team dat het experiment uitvoert. Dat kan leiden tot zogenaamde cross-over effecten tussen de treatments welke worden versterkt door de extra status die deelname aan het experiment en het onderzoek met zich meebrengt en de bijdrage aan het onderzoek zelf (o.a. invulling van procesevaluatie vragenlijsten). In de RCT-experimenten literatuur spreekt men dan over gedragseffecten in dit geval bij de begeleiders die een zuivere vergelijking tussen de vier experimentgroepen bemoeilijkt. De deelname aan het experiment door de deelnemers en de speciale behandeling die ook zij daarbij krijgen zoals een uitnodiging bij de start van het experiment aan bijeenkomsten met projectleider en coach, het contact met coaches over het invullen van de vragenlijsten (de deelnemers worden enkele keren eraan herinnerd de vragenlijst in te vullen en hebben dan bijvoorbeeld telefonisch contact met de coaches) kunnen leiden tot gedragseffecten bij de deelnemers waardoor zij zich anders gedragen dan bijstandsgerechtigden die niet aan het experiment deelnemen. Dit wordt in de literatuur het 'Hawthorne'

effect genoemd³. Dit heeft uiteraard in dat geval invloed op het onderzoek naar de effecten van het experiment. Mede om die reden is ook nog een referentiegroep onderscheiden waarin alle tot de populatie behorende bijstandsgerechtigden die niet aan het experiment deelnemen zijn vertegenwoordigd. Deze worden administratief gevolgd waardoor inzicht verkregen wordt in de mate waarin het experiment bijdraagt aan een grotere uitstroom naar betaald werk. De onderzoekers zullen uiteraard in het onderzoek met deze mogelijke verandering in de standaardtreatment sinds de start van het experiment per 1 oktober 2017 rekening houden. Van belang is dat ook de overige ‘treatments’ zoveel mogelijk in stand worden gehouden en niet veranderen maar ook dat is niet vanzelfsprekend aangezien beleid voortdurend verandert en het experiment invloed heeft op de begeleiding door de klantmanagers in de treatmentgroepen. In de praktijk blijkt dit een lastige opgave omdat er experiment -en leereffecten optreden bij de klantmanager waardoor de treatment zich wijzigt. De vergelijking met de referentiegroep (wat uitstroom naar werk betreft) is daardoor erg belangrijk vooropgesteld dat de treatmentgroepen en de referentiegroepen wat samenstelling betreft vergelijkbaar zijn. Deze vergelijking met de referentiegroep is overigens alleen mogelijk bij de uitkomstmaat uitstroom naar werk, maar niet bij de overige uitkomstmaten van welbevinden, gezondheid, zelfredzaamheid, keuzemogelijkheden en sociale participatie. Daarvoor ontbreekt een referentiegroep omdat over deze uitkomstmaten geen administratieve data beschikbaar zijn.

2.2 Procedure en werving

Voor het onderzoek is er gebruik gemaakt van een RCT design, dat wil zeggen een gerandomiseerd experiment met een controlegroep⁴, waarbij deelnemers aselekt worden toegewezen aan de ‘treatments’ of behandelingsgroepen. De doelgroep omvatte alle bijstandsgerechtigden in Apeldoorn/Epe met uitzondering van enkele specifieke groepen die zijn uitgesloten van deelname aan het experiment omdat ze niet voldeden aan de uitsluitingscriteria. Uitgesloten zijn mensen die geen recht (meer) hebben op een uitkering Participatiewet, mensen die in een inrichting verblijven, dak -en thuislozen, mensen die gedurende het experiment met pensioen gaan (>65 jaar), mensen die de Nederlandse taal niet machtig zijn, mensen die volledig arbeidsongeschikt zijn en mensen die andere regelingen zitten (IOAW, IOAZ, BBZ) bij aanvang van het experiment.

Steekproef en werving

In Apeldoorn en Epe is de steekproef voor de helft getrokken uit het zittende bestand en voor de helft uit de “nieuwe instroom”, dat zijn alle nieuwe bijstandsgerechtigden die vanaf 1 juli 2016 zijn geregistreerd als mensen met een lopende uitkering. Er is gekozen voor een gestratificeerde steekproef om er zeker van te zijn dat er voldoende deelnemers uit het “zittende bestand” vertegenwoordigd zijn in het experiment. De bevindingen uit de literatuur waren immers dat het grootste effect van de andere

³ Een variant hierop is het “John Henry” effect waarbij er competitie (“rivalry”) ontstaat tussen de treatments en de controlegroep waarbij klantmanagers/consulenten in de controlegroep zich extra inzetten voor bijvoorbeeld uitstroom naar werk om niet onder te doen voor de collega’s of wanneer deelnemers zich extra gemotiveerd tonen voor deelname aan het experiment, ook in de controlegroep, om de begeleiding succesvol te laten verlopen.

⁴ Engels: Randomised Controlled Trial (RCT).

bejegening juist verwacht mocht worden bij deze groep (zie Muffels & Klink, 2017). Uit de bijstandspraktijk is immers bekend dat na verloop van tijd het aantal contacten met de cliënten met mensen uit het 'zittende bestand' gering is. Juist deze cliënten blijken baat te hebben bij 'extra aandacht'. De datum van 1 juli 2016 is gekozen om het beoogde aantal van 450 deelnemers in Apeldoorn en 90 in Epe (zie bijlage 1) binnen 6 maanden na de start van het experiment per 1 oktober 2017 te hebben bereikt. In eerste instantie zijn 1125 mensen in Apeldoorn en 225 mensen in Epe in de steekproef opgenomen, later is de steekproef uitgebreid tot alle mensen met een lopende uitkering met uitzondering van bepaalde cliëntgroepen zoals arbeidsongeschikten in de WIA/Wajong, mensen met ernstige fysieke of mentale gezondheidsproblemen, ouderen boven 65 jaar en statushouders die het Nederlands niet machtig zijn. Uiteindelijk is niet meer dan 18% van de totale populatie uitgesloten van deelname zodat de totale steekproef uiteindelijk bestond uit 4191 mensen. Het streefgetal van 150 respectievelijk 30 deelnemers per groep, dus 450 in totaal in Apeldoorn en 90 in Epe, is voor wat betreft Apeldoorn gebaseerd op een poweranalyse die in maart 2017 is uitgevoerd. Daarbij is uitgegaan van een power van 95%. Bij Epe is er op statistische gronden van uitgegaan dat een aantal van 30 deelnemers voldoende is om de specifieke resultaten voor Epe op een betrouwbare manier te kunnen schatten. Daartoe worden voor de analyse van de effecten de gegevens van Epe en Apeldoorn samengevoegd.

De verwachting was dat bij een streefgetal van 540 deelnemers (450 Apeldoorn, 90 Epe) en dus 270 nieuwe instromers (225 in Apeldoorn en 45 in Epe) er 6-9 maanden van werving van deelnemers nodig zijn om dit streefgetal te bereiken. Dat is niet helemaal gelukt. De laatste instromer dateert van 25 september 2018. Deze deelnemer kan dus slechts 12 maanden meedoen aan het experiment terwijl de deelnemers uit het zittende bestand die op 1 oktober 2017 zijn ingestroomd de volle twee jaar deelnemen. De gemiddelde duur van deelname aan het experiment in Apeldoorn en Epe is overigens 1,9 jaar. Op 1 juni 2017 bestond de totale doelpopulatie, dus voor toepassing van de uitsluitingscriteria, uit 3750 bijstandsgerechtigden in Apeldoorn en 441 bijstandsgerechtigden in Epe met een lopende uitkering in de Participatiewet. Door toepassing van de uitsluitingscriteria werden 18% van de bijstandsgerechtigden in Apeldoorn en 21% van de bijstandsgerechtigden in Epe niet toegelaten tot het experiment. Dat resulteerde in een netto aantal van 3438 bijstandsgerechtigden in de steekproef, 3091 in Apeldoorn en 348 in Epe.

In totaal hebben zich tot 25 september 2018, de dag waarop de laatste deelnemer is geworven, 786 mensen aangemeld (659 in Apeldoorn en 127 in Epe) waarvan 497 in Apeldoorn en 90 in Epe voldeden aan de toelatingscriteria. Daarvan hebben 96 potentiële deelnemers, 87 in Apeldoorn en 9 in Epe, zich weer afgemeld. Deze 96 afmelders deden zich in twee van de drie gevallen voor bij het zittende bestand. Dat zijn alle mensen met een uitkering die is gestart voor 1 oktober 2017 waardoor er netto 491 deelnemers zijn, 410 in Apeldoorn en 81 in Epe⁵. Daarvan is 57% afkomstig uit het zittende bestand per 1 juni 2016 maar 80% uit het zittende bestand per 1 oktober 2017.

⁵ In de tussenrapportage no. 3 van 1 Maart 2019 werd nog een netto aantal van 470 deelnemers genoemd, dat is exclusief afmeldingen. Door een nieuwe kritische beoordeling van de 'echte' afmeldingen is dit aantal in deze rapportage bijgesteld naar boven.

Afmeldingen

De deelnemers worden aselekt toegewezen aan de drie groepen, maar deelnemers mogen na indeling alsnog weigeren deel te nemen aan het experiment. Ze kunnen dan echter niet worden geplaatst in een andere groep. Uiteindelijk zijn er 96 'echte' afmeldingen geregistreerd⁶ in Apeldoorn (n=87)/Epe(n=9) hetgeen 12% is van het aantal aanmeldingen (N=787). De meeste afmeldingen (twee van de drie) komen uit het zittende bestand per 1 oktober 2017. Steeds is door de gemeente de reden van afmelding geregistreerd. Nadere analyse van de 96 afmeldingen leerde dat 29 (30%) daarvan plaatsvonden voor de start. Ze zagen ofwel af van deelname voor de start of hun uitkering was beëindigd door verandering in de situatie (verhuizing, aangaan relatie) of doordat het recht op uitkering was herzien door fraude of niet voldoen aan andere verplichtingen. Twee deelnemers hebben zich afgemeld vanwege de indeling in de groep, 3 deelnemers omdat met het onderzoek te belastend vond en 40 deelnemers zagen het nut van deelname niet (meer) in (bij elkaar 47%). De andere afmeldingen zijn ontstaan doordat men taalproblemen of ernstige sociale problemen had of kampte met een slechte fysieke en/of mentale gezondheid (n=7 ofwel 7%). Het resterende deel is uitgestroomd vanwege verhuizing, aangaan relatie, inkomens- vermogenstoets of overgang andere uitkering (10%) of vanwege uitsluiting door de consulent vanwege een overvolle groep (5%).

2.3 Onderzoekpopulatie: deelnemers in de verschillende treatments

In Tabel 2.3.1 wordt weergegeven in hoeverre de afmeldingen al dan niet gelijkelijk zijn verdeeld over de treatments. Tevens geeft de tabel de bereikte aantallen weer voor elk van de drie groepen inclusief en exclusief de 96 afmelders. Tot slot wordt in de tabel de bruto en netto response gegeven in procenten. De verwachting was dat het meer moeite kost om mensen in de standaardgroep te krijgen en te behouden aangezien zij in beginsel weinig baat hebben bij deelname. Dit bleek uiteindelijk erg mee te vallen. De netto aantallen, dus na aftrek van de afmeldingen, zijn redelijk tot goed gespreid over de drie interventiegroepen. De eigen regiegroep in Apeldoorn en in Epe blijft wat achter met 128 en 22 netto deelnemers respectievelijk tegenover 142 en 30 deelnemers voor de beide andere groepen. De bruto response is 86% in Apeldoorn en 91% in Epe op basis van de 'echte' afmeldingen. De netto response is 74 en 82% respectievelijk. Dit zijn hoge response percentages welke op het conto van de gemeente geschoven kunnen worden die erg veel moeite hebben gedaan de deelnemers te motiveren en te overtuigen de vragenlijsten in te vullen. Het succes van de werving voor het experiment kan voor een belangrijk deel verklaard worden door de persoonlijke motivatie, enthousiasme en inzet van de medewerkers van de gemeente (projectleiders, databasemanager, ondersteunende beleidsmedewerkers, klantmanagers, afdelingshoofden, politieke steun vanuit college i.c. de wethouder).

⁶ Het totaal aantal aanmeldingen in Apeldoorn en Epe was 787 waarvan 199 niet voldeden aan de toelatingscriteria of waarvan de uitkering reeds voor de start van het experiment was beëindigd.

Tabel 2.3.1: Deelnemers, afmeldingen in Apeldoorn en Epe per 'treatment groep' plus response percentages, per 1 oktober 2019 (n==580)¹⁾

Treatments/Gemeente	Apd		Epe		Apd	Epe	Totaal
Treatments/Afmeldingen	Deeln.	Afm.	Deeln.	Afm.	Netto	Netto	Netto
Zelf aan het stuur (eigen regie)	161	33	27	5	128	22	150
Op maat (intensieve begeleiding)	170	28	32	2	142	30	172
Standaardgroep	159	18	31	2	141	29	170
Totaal	490	79 ²⁾	90	9 ²⁾	411	81	492
Bruto en netto response (in %)	86%	14%	91%	9%	74%	82%	75%

Noot:¹⁾ ²⁾ Van 6 afmeldingen in Apeldoorn en 2 in Epe is de indeling in treatment niet bekend

Bron: Gegevens gemeente Apeldoorn/Epe, Tilburg University controlebestand en BUS-gegevens, 1 oktober 2019

Wat goed heeft gewerkt is een intensieve persoonlijke benadering waarin de cliënt persoonlijk werd geënthousiasmeerd voor deelname hetgeen vanuit het gezichtspunt van het aanwezige 'wantrouwen' van de cliënt ten opzichte van de gemeente geen sinecure is. Deze persoonlijke benadering was ook nodig om deelnemers te stimuleren de online vragenlijst voor de metingen in te vullen. Dat deelnemers in de periode dat het experiment liep gemakkelijker uitstromen is vooral te danken aan de economisch gunstige situatie. Voor een belangrijk deel zijn het vooral de jongeren met startkwalificatie die daarvan profiteren en niet degenen zonder startkwalificatie en de ouderen.

Randomisatie en representativiteit

Doordat de deelnemers aselekt zijn toegewezen aan de drie interventiegroepen verwachten we geen grote verschillen tussen de deelnemers en de doelpopulatie waaruit de steekproef is getrokken. Mensen moeten zich evenwel aanmelden waardoor er zelfselectie zou kunnen ontstaan. Daarnaast kan zelfselectie ontstaan doordat mensen met bepaalde kenmerken zich na de aanmelding alsnog afmelden dan andere. Het zou kunnen zijn dat lager opgeleiden of ouderen boven 55 jaar doordat ze meer moeite hebben met invulling van online vragenlijsten ondervertegenwoordigd zijn. Voorts verwachten we dat mensen met weinig kansen op de arbeidsmarkt minder genegen zijn zich aan te melden, veelal betreft dat mensen die reeds lang in de bijstand zitten en tot het zogenaamde 'zittende bestand' behoren. In de volgende drie tabellen worden de deelnemers vergeleken met de referentiegroep (dat is de populatie van bijstandsgerechtigden die voldoet aan de toelatingscriteria minus degenen die zich hebben aangemeld als deelnemer). Daarbij hebben we gebruik gemaakt van de BUS gegevens op individueel niveau bij de start van het experiment. Voor het zittende bestand zijn dat gegevens per 1 oktober 2017. Voor de nieuwe instromers zijn dat de gegevens ultimo van de maand die zo dicht mogelijk ligt bij de start van het experiment veelal de maand ervoor. De informatie beschikbaar in de BUS betreft gegevens zoals geslacht, leeftijd, hoogst afgeronde opleidingsniveau, duur uitkering, nationaliteit, leefvorm en huisvestingssituatie.

Tabel 2.3.2 Representativiteit onderzoekspopulatie (inclusief afmelder) naar achtergrondkenmerken bij start van experiment, Apeldoorn/Epe, 2017-2018 (in %)

		-1-	-2-	-3-	-4-	-5-
Achtergrondkenmerken bij start (BUS)	Categorie	Deelnemers	Eigen regie	Intensieve begeleiding	Standaardgroep	Referentiegroep
Geslacht	Man	42.6	42.9	45	39.6	45.2
	Vrouw	57.4	57.2	55	60.4	54.8
Leeftijd	gemiddeld	45.3	46.5	44.9	44.5	42.4
	<27 jaar	8.6	7.7	9.0	9.1	14.7
	27-34 jaar	17.6	17.5	15.0	20.4	16.2
	35-44 jaar	24.1	16.9	29.5	25.3	22.9
	45-54 jaar	28.3	31.7	29.0	24.2	24.1
	55 jaar (tot 65)	21.4	26.2	17.5	21.0	22
Opleidingsniveau	Laag	54.6	56.5	56.4	50.8	57.6
	Midden	23.9	23.7	22.8	25.4	16.6
	Hoog	6.2	7.5	5.5	5.8	4.4
	Geen opl./onbekend	15.3	12.4	15.4	18.0	21.5
Migratieachtergrond	Buitenlands	5.7	7	5.4	4.8	24
	Nederlands	94.3	93	94.6	95.2	76
Huisvesting situatie	Eigenaar	3.6	3.3	3.5	4.3	1.8
	Huurder	87.4	90.8	87.5	86.6	87.2
	Inwonend	7.3	5.4	7.5	9.1	8.9
	Overig	1	1	1.5	0	2.2
Leefvorm	Alleenstaande ouder	22.2	21.7	23.5	21.4	15.5
	Alleenstaande	58.7	58.7	58.5	58.8	56.7
	Samenwonend	19.1	19.6	18	19.8	27.8
Duur uitkering	Tot 1 jaar	23.0	22.5	23.4	23.2	13.2
	1 tot 3 jaar	32.6	27.5	36.6	33.5	32.6
	3 jaar of meer	44.3	50	40.1	43.2	54.2
	Gemiddelde duur	2.2	2.3	2.0	2.1	3.3
Aantal (N=)	Totaal	571	184	200	187	3822

Bron: BUS-gegevens Apeldoorn/Epe, 1 oktober 2019

Noot: Achtergrondkenmerken van deelnemers in de onderzoeksgroep (kolom 4) en in de twee experimentgroepen (kolom 2-3). T1: Eigen regie-Zelf aan het stuur; T2: Intensieve begeleiding-Op maat; T3. Controlegroep=Standaardgroep; Referentiegroep (kolom 5: niet-deelnemers, wel in doelpopulatie). Het aantal van 571 deelnemers betreft deelnemers plus afmeldingen, dus iedereen die zich heeft aangemeld en tot de doelpopulatie behoorde en waarvoor informatie in de BUS-data beschikbaar is. Van 9 deelnemers ontbrak dus informatie in de BUS. De informatie is gemeten bij de start van deelname aan het experiment.

De referentiegroep omvat 3822 mensen waartoe mensen behoren met een lopende uitkering per 1 oktober 2017 of mensen die tussen 1 oktober 2017 en 1 juli 2018 zijn ingestroomd in de bijstand maar hebben besloten niet deel te nemen aan het experiment. De drie treatment groepen betreffen alle 580 mensen die zich hebben aangemeld voor het experiment inclusief de afmeldingen (492 deelnemers plus 88 afmeldingen). Voor 9 aanmeldingen ontbraken BUS gegevens. Nadere analyse (zie Tabel 2.3.2) laat zien dat er slechts kleine verschillen zijn tussen de twee treatment groepen en de standaardgroep. Vanwege het RCT design zijn deelnemers aselect toegedeeld aan deze drie groepen juist om de groepen qua samenstelling zoveel mogelijk vergelijkbaar te maken. Dat wordt randomisatie genoemd. Deze lijkt dus goed gelukt.

Selectie bij aanmelding

Door selectie (toelatingscriteria) en zelfselectie in het experiment (wie melden zich aan en wie melden zich af) zijn er evenwel iets grotere verschillen ontstaan tussen de experimentgroepen en de referentiegroep, vooral naar nationaliteit en enigszins bij duur van de uitkering, maar deze verschillen naar duur vallen bij nader inzien erg mee. Kleine verschillen zijn daarnaast ontstaan in verband met leeftijd (in de drie treatmentgroepen zijn jongeren iets oververtegenwoordigd), vrouwen zijn licht oververtegenwoordigd evenals bijstandsgerechtigden met een middelbare opleiding.

Naar duur is er een lichte oververtegenwoordiging van korte duren en een ondervertegenwoordiging van lange duren (>3 jaar). De bijstandsgerechtigden met duren van 1-3 jaar blijken daarentegen evenredig te zijn vertegenwoordigd in de treatment groepen. Dit betekent dat een vergelijking van de uitkomstmaten en zeker ook de uitstroomkansen naar werk tussen de vier groepen die sterk afhankelijk is van de duur van de uitkering, rekening moet houden met deze lichte onder- en oververtegenwoordiging naar duur van de uitkering en nationaliteit voorafgaand aan het experiment. Dit hebben we gedaan met behulp van logistische regressie en matching technieken. In de logistische regressie voor de uitstroomkans naar betaald werk corrigeren we voor verschillen in duur door de duur bij de start van het experiment op te nemen in de regressievergelijking. In de matching procedure gaan we op zoek naar mensen in de bijstandspopulatie van niet-deelnemers die zoveel mogelijk lijken op mensen die aan het onderzoek deelnemen en dus ook qua duur uitkering, leeftijd, geslacht, nationaliteit en dergelijke vergelijkbaar zijn.

2.4 Deelnemers in de product- en procesevaluatie

Het voorliggende onderzoek bestaat zoals eerder aangegeven uit twee onderdelen: de product –en procesevaluatie (zie Muffels & van der Klink 2017). De *productevaluatie* is erop gericht na te gaan wat beter werkt voor de cliënt; de nieuwe regelluwe aanpak of de standaardaanpak met een strenge toetsing van verplichtingen en de toepassing van maatregelen en kortingen. Over een periode van 2 jaar zijn deelnemers aan het experiment gevolgd. In het onderzoek wordt niet alleen gekeken naar of mensen eerder betaald werk vinden maar ook wat het met mensen doet, met hun vertrouwen, gezondheid, hun welbevinden, hun zelfvertrouwen en hun zelfredzaamheid. Daartoe vulden deelnemers aan het begin van het experiment, na 1 jaar en na 2 jaar een online vragenlijst in. De eerste noemen we de nulmeting en de laatste de eindmeting. De uitkomsten van de analyses op deze vragenlijsten zullen een antwoord geven op eerder gestelde onderzoeksvragen. Wanneer deelnemers uitstromen naar werk of bijvoorbeeld verhuizen, blijven ze deelnemer en worden ze gevraagd de vragenlijsten in te vullen. De bereidheid om

dat te doen zal evenwel naar verwachting geringer zijn bij de uitstromers waardoor de non-response naar verwachting zal toenemen naarmate meer mensen uitstromen. In het onderzoek worden degenen die zich hebben afgemeld na de start van het experiment ook na afmelding administratief gevolgd. Dat gebeurt via de gemeentelijke administratie (BUS) alsook door koppeling van de gegevens van deze oorspronkelijke deelnemers aan de bestanden van het CBS. Dit is voor een valide en zuivere meting van de uitkomsten van het RCT experiment noodzakelijk. De wijze van gegevensverzameling is getoetst aan de AVG privacy regels en aan de ethische procedures voor 'informed consent' welke gelden voor wetenschappelijk onderzoek. De persoonsgegevens (NAW) zijn door de gemeente geanonimiseerd met behulp van een anoniem identificatienummer waarbij alleen de gemeente en het CBS toegang tot de persoonsgegevens hebben. Daartoe is een bewerkingsovereenkomst met de gemeente gesloten.

De *procesevaluatie* is erop gericht na te gaan hoe de projectleider(s) en klantmanagers/ (werk)coaches als professional het proces van implementatie van het experiment in de gemeente percipiëren en waarderen, en welke meer -of minderwaarde de klantmanagers/coaches ervaren van de andere bejegening van de cliënt voor de uitoefening van hun taken in de nieuwe aanpak. Daartoe zijn er regelmatig focusgroep gesprekken gehouden met de projectleider en de coaches. In totaal zijn er 4 focusgroep gesprekken gehouden. Het eerste gesprek vond plaats op 6 december 2017 nadat het experiment enkele maanden van start was gegaan en het eerste gesprek met de cliënt is gehouden. Het tweede gesprek heeft op 3 oktober 2018 plaatsgevonden, het derde gesprek volgde op 12 juni 2019 en het vierde gesprek op 30 oktober 2019. Ten tweede vulde de coach aan het begin van het experiment, na 1 jaar en na 2 jaar een online vragenlijst in over de afspraken en bevindingen van de coaching gesprekken. Tezamen met het eerste focusgroep gesprek vormt dit de nulmeting van het procesevaluatie-deel. Ten slotte hebben de klantmanagers aan het einde van het experiment (november 2019) een afsluitende vragenlijst ingevuld over hun eigen ervaringen gedurende het experiment. Door het uitvoeren van deze procesevaluatie kan inzicht worden verkregen in de effectiviteit van de interventie, maar ook de mate waarin ineffectiviteit is terug te voeren tot een 'theory failure' (beleidstheorie) of een 'implementation failure' (uitvoering). In de volgende paragrafen zal de dataverzameling van zowel de product- als procesevaluatie worden toegelicht.

2.5 Vragenlijsten product- en procesevaluatie

2.5.1 Productevaluatie: vragenlijsten

De vragenlijsten bevatten vragen over de onderwerpen gezondheid en welbevinden, zelfredzaamheid, zoek- en sollicitatiegedrag, ervaren keuzemogelijkheden (capabilities), sociale participatie, sociaal vertrouwen, vertrouwen in overheid en financiële situatie en armoede. Wanneer deelnemers uitstromen, bijvoorbeeld naar werk, blijven ze deelnemer en worden ze gevraagd de vragenlijsten in te vullen. In de praktijk blijkt dat ze zich dan sneller afmelden. Het gevolg hiervan is dat de respons afneemt naarmate het experiment vordert. Het doel van de eerste vragenlijst (de zogenaamde nulmeting) was om de situatie van de cliënt vast te stellen aan het begin van het experiment. Het invullen ervan gebeurde tijdens of direct na het eerste gesprek met de deelnemer aan het experiment. De tweede en derde vragenlijst (ook wel eerste en tweede vervolgmeting) zijn 1 jaar en 2 jaar na de start van de begeleiding verstuurd. Omdat er tot 1 juli 2018 nieuwe deelnemers zijn ingestroomd in het experiment, gebeurde het versturen van de

vervolgvragenlijsten in tranches. Hierbij werden in drie tranches de deelnemers in volgorde van aanmelding geselecteerd voor invulling van de tweede en derde vragenlijst. Er is gekozen voor twee vervolgvragenlijsten om effecten over de tijd heen in beeld te kunnen brengen. Deelnemers in de referentiegroep zijn niet benaderd voor een vragenlijst.

De deelname aan het experiment is zoals eerder opgemerkt vrijwillig, men mag te allen tijde het experiment weer verlaten maar er wordt wel van een deelnemer verwacht dat deze meedoet aan het onderzoek en de vragenlijsten zorgvuldig invult. Daarover zijn afspraken gemaakt met de klant, hoewel het in de praktijk soms moeilijk is gebleken om een deelnemer aan deze afspraak te houden.

Invulling deelnemer vragenlijst (productevaluatie)

Een persoonlijke benadering bleek uitermate belangrijk om deelnemers te stimuleren de online vragenlijst voor de nulmeting in te vullen. Zoals hiervoor toegelicht heeft dat veel energie gevraagd. Het vraagt overredingskracht en uitleg van de klantmanager om de deelnemer te motiveren en te overtuigen de vragenlijst in te vullen. Vanaf de start van het experiment op 1 oktober 2017 jl. hebben de deelnemers aan het experiment de online vragenlijst ingevuld. De bereidheid bij de deelnemers om de tweede en vooral ook de derde vragenlijst in te vullen bleek sterk te zijn afgenomen na de nulmeting ondanks een intensieve campagne gericht op invulling door de deelnemer opgezet door de gemeente Apeldoorn/Epe. Veel vragen gaan over werk zoeken en vinden hetgeen voor de mensen die na 2 jaar nog geen werk hebben gevonden ondanks de andere bejegening en begeleiding geen haalbare kaart is waardoor de motivatie om de tweede en derde vragenlijst in te vullen veel geringer bleek. De gemeente Apeldoorn/Epe heeft met het oog op het belang voor het onderzoek van een goede invulling zich ingespannd om mensen te overreden de vragenlijst in te vullen. Dat heeft uiteindelijk geresulteerd in een respons van ongeveer 54% voor de beide laatste deelnemers vragenlijsten tegenover 93% in de nulmeting. Uiteindelijk is een respons bereikt van 67% hetgeen bij deze als ‘moeilijk bereikbare’ onderzoeksgroep bekend staande groep als ‘goed’ mag worden bestempeld zeker indien in beschouwing wordt genomen dat ruim 35% van de deelnemers na 2 jaar om vele redenen is uitgestroomd uit de bijstand vanwege verhuizing, herzien recht op uitkering, overgang naar andere uitkeringen maar ook door het vinden van werk. In totaal heeft 21% in de periode van 2 jaar werk gevonden (inclusief uitstroom naar zelfstandigheid) en heeft 14% van de deelnemers de uitkering beëindigd om andere redenen. De 1065 door hen “online” ingevulde vragenlijsten⁷ zijn door de onderzoekers nader geanalyseerd. Voordat we de resultaten verder bespreken wordt hieronder inzicht geven in de bereikte aantallen ingevulde vragenlijsten per treatment groep. Tevens wordt voor iedere meting het response percentage vermeld. Het responsepercentage neemt over de tijd sterk af door uitstroom. Na uitstroom zijn vanwege bijvoorbeeld verhuizing of uitstroom naar werk of studie deelnemers veel minder genegen om de vragenlijst alsnog in te vullen. In meting 2 en 3 valt het response percentage terug van 95 naar 54%. De uitstroom in de 2-jarige periode is 35%.

Procesevaluatie: klantmanagers vragenlijsten

Ten behoeve van de procesevaluatie zijn er ook periodiek vragenlijsten ingevuld door de klantmanagers die de deelnemers in begeleiding hebben. Voor elke deelnemer die de klantmanager in begeleiding heeft, vult een klantmanager een vragenlijst in. Net als bij de deelnemers vragenlijsten zijn deze klantmanager

⁷ De deelnemers hebben de vragenlijst online ingevuld. De deelnemers kregen een door de Universiteit gemaakte persoonlijke weblink opgestuurd waarmee zij de vragenlijst in de thuissituatie online konden invullen.

vragenlijsten drie keer afgenomen: aan het begin, na 1 jaar en na 2 jaar begeleiding (maar wel in volgorde van aanmelding van een klant). In deze vragenlijsten zijn vragen gesteld over het proces van begeleiding, zoals de doelen die zijn ingezet in overleg met de klant, de activiteiten die zijn ondernomen en de trajecten die zijn aangevraagd, de motivatie van de deelnemer voor de begeleiding door de klantmanager, een inschatting van de klantmanager over de kansen op de arbeidsmarkt van de deelnemer en de leereffecten van het experiment voor zowel de klantmanager als de klant. Wanneer een deelnemer uitstroomt, bijvoorbeeld vanwege het vinden van betaald werk of vanwege verhuizing, dan is bij ontbreken van contact het lastig voor de klantmanager om de vragenlijst in te vullen.

Tabel 2.5.1: Aantallen ingevulde vragenlijsten voor de deelnemers en klantmanagers, Apeldoorn/Epe, 2017-2019 (N=551)¹⁾

	Deelnemers					Klantmanagers/consulenten		
Treatments/Groepen	CV1	CV2	CV3	Totaal	PCEV1	PCEV2	PCEV3	Totaal
Eigen Regie	153	78	85	316	151	146	71	368
Intensieve Begeleiding	188	112	107	407	181	176	126	483
Standaardgroep	174	91	77	342	182	196	134	512
Totaal meting 1,2,3 (N=551) (N=522)(N=513)	515	281	269	1065	514	518	331	1363
Bruto response %	93%	54%	52%	67%	93%	99%	65%	85%

Noot: CV=Cliënt vragenlijsten; PCEV=Procesevaluatievragenlijsten. De response percentages zijn berekend uitgaande van het bruto aantal deelnemers in de verschillende metingen (580 aanmeldingen minus 29 afmeldingen voor de start). Tussentijds is de verhouding tussen aanmeldingen en afmeldingen aangepast doordat een aantal aanmeldingen en afmeldingen niet aan de criteria voldeden.

Bron: Gegevens vragenlijsten, Tilburg University, 1 oktober 2019

Voor vier van de vijf deelnemers zijn desondanks de klantmanagers in staat gebleken de vragenlijst in te vullen. Dit is een goed resultaat en te danken aan de proactieve houding en inzet van de klantmanagers, projectleiders en management. In totaal zijn 1363 vragenlijsten ingevuld, hetgeen een respons percentage betekent van 85%⁸. In Tabel 2.5.1 wordt een overzicht gegeven van de bereikte aantallen ingevulde vragenlijsten voor de procesevaluatie alsmede de response percentages voor de verschillende metingen.

2.5.2 Administratieve BUS data

Zoals opgemerkt zullen in mei 2020 de gegevens geupload worden naar het CBS om ook vergelijkingen met andere gemeenten mogelijk te maken. In deze rapportage maken we gebruik van de gemeentelijke BUS bestanden om bijvoorbeeld de, voorover deze aanwezig zijn, uitstroomkansen naar werk te bepalen maar ook om te corrigeren voor compositieverschillen tussen de treatment groepen voor zover deze aanwezig zijn. Deze bestanden bevatten informatie over demografische kenmerken van de bijstandspopulatie, de reden van beëindiging van de uitkering vanwege uitstroom uit de uitkering, de arbeidsgeschiedenis, de loonwaarde, het aantal en type contactmomenten en de ondernomen re-

⁸ Wordt in de derde meting van de procesevaluatie ook gecorrigeerd voor uitstroom dan bedraagt het totale response percentage 92%.

integratieactiviteiten afkomstig uit de registratiesystemen van de gemeente Apeldoorn/Epe. Daarnaast worden de data geupload naar het Centraal Bureau voor de Statistiek (CBS) om deze te verrijken met gegevens over werk (loopbaan) en inkomen, uitkeringen en andere achtergrondkenmerken.⁹ Hierdoor is het mogelijk om de loopbanen van de deelnemers te reconstrueren en ook gedeeltelijke uitstroom naar flexibele banen, (kleine) deeltijdbanen en vaste banen te onderzoeken. Door het gebruik van CBS-data zijn de gegevens tussen de experimenten met de Participatiewet bovendien beter vergelijkbaar.

3 Resultaten product -en procesevaluatie

3.1 Productevaluatie en uitkomstmaten

In de productevaluatie gaat het niet alleen om de uitkomstmaat uitstroom naar betaald werk maar ook om de verbetering van de gezondheid, het welbevinden, het zelfvertrouwen, de keuzemogelijkheden, de sociale participatie en de zelfredzaamheid van de cliënt. Eerst bespreken we de resultaten van de analyse op de uitkomstmaat uitstroom naar betaald werk ofwel gedeeltelijke uitstroom ofwel volledige uitstroom. Onder volledige uitstroom verstaan we dat bijstandsgerechtigden vanwege uitstroom naar betaald werk volledig onafhankelijk worden van de uitkering. Bij een loon ter hoogte van het minimumloon betekent dit dat men minimaal 27 uur gaat werken (70% van 38-40 uur) aangezien de uitkering voor een alleenstaande 70% bedraagt van het volwassen minimumloon.

Deeltijdse uitstroom en uitstroom naar tijdelijk werk

Er zijn in de BUS data uitsluitend gegevens beschikbaar over volledige uitstroom uit de uitkering maar niet over deeltijduitstroom of over uitstroom naar tijdelijk werk. De gemeenten beschikken over uitkeringsgegevens van mensen in de bijstand inclusief de korting op de uitkering vanwege bijverdiensten op grond waarvan een goede inschatting gemaakt zou kunnen worden van de aantallen mensen die deeltijds in plaats van voltijds uitstromen en voor hoeveel uren. Deze info was voornamelijk nog niet beschikbaar. In plaats daarvan zal gebruik worden gemaakt van de gegevens in de CBS microdatabestanden. De data van Apeldoorn/Epe zullen daartoe begin Maart 2020 worden geupload naar de CBS omgeving. De tweede en derde cliëntvragenlijst bevatten wel enige informatie van de deelnemer over zijn of haar situatie op het moment van invulling van de vragenlijst (zie Tabel 3.1.1). Deze geeft evenwel geen volledig beeld van de uitstroom naar deeltijd -of tijdelijk werk. Op basis van de vragenlijsten afgenomen bij de deelnemers is bekend of ze bij meting 2 en 3, dus 1 en 2 jaar na de start, een voltijdse, deeltijdse of tijdelijke baan hebben of dat ze andere activiteiten hebben ondernomen zoals vrijwilligerswerk, een opleiding of studie of een baan als freelancer of zelfstandige. We weten niet of dat een nieuwe baan sinds de start van het experiment is of een baan die ze ook al hadden bij de start.

⁹ In alle gebruikte databestanden werden persoonsgegevens bij voorbaat door middel van een sleutel vervangen door een betekenisloos administratienummer (pseudonimiseren). Middels dit nummer kunnen databestanden uit verschillende bronnen en over de tijd heen aan elkaar worden gekoppeld. Omdat geen persoon toegang heeft tot zowel de databestanden als de sleutel, is informatie niet naar deelnemers herleidbaar.

Tabel 3.1.1: Type sociale participatie activiteiten in metingen 2 en 3, Apeldoorn/Epe

Type activiteit	CV-Meting 2+3	
	N=	%
Voltijdse baan	31	6%
Deeltijdse baan, uitzendwerk, enkele uren werk	78	14%
Zelfstandige	3	1%
Werkstage, proefplaatsing, overig werk (voorbereiding werk)	8	1%
Opleiding, cursus of studie	13	2%
Gedeeltelijke/volledige ao-uitkering	26	5%
Sociale werkvoorziening	18	3%
Vrijwilligerswerk/mantelzorg	137	25%
Bijstand	219	39%
Verhuisd	5	1%
Ziek, overig	4	1%
Wn/ga	16	3%
Totaal N=	558	100%

Bron: Vragenlijsten eerste en tweede vervolgmeting deelnemers , TiU, 2019

Uit de gegevens in tabel 3.1.1 blijkt dat in de beide metingen 10% volledig is uitgestroomd als werknemer (inclusief sociale werkvoorziening) of als zelfstandige, 1,5% werkt in een proefplaatsing of stage. Opvallend is het grote aantal mensen dat zegt in deeltijd te werken (n=78); dat betekent dat 14% van alle deelnemers een aantal uren per week in een kleine baan werkt, in een uitzendbaan of een deeltijdbaan. Dat betekent dat ongeveer een op de vier deelnemers naar eigen zeggen volledig of gedeeltelijk aan het werk is gegaan in de tweejarige experimenteer periode. Opvallend is ook het relatief grote aantal deelnemers (25%) dat zegt vrijwilligerswerk te verrichten. De gunstige arbeidsmarktsituatie in combinatie met de inspanningen van de gemeente voor het experiment zullen hier zeker debet aan zijn. Voor twee op de drie deelnemers die werk vindt is dit een deeltijdse baan of uitzendwerk. In termen van sociale participatie is ongeveer de helft van de deelnemers actief bezig met een deeltijdse of voltijdse baan (25%) of vrijwilligers werk (25%).

Type baan sinds vorige keer

Daarnaast is gevraagd of ze sinds de vorige keer (meting) een baan hebben gevonden en wat voor type baan dat is. De respons in de tweede en derde cliëntvragenlijst bedroeg evenwel slechts 60% waardoor geen volledig beeld mogelijk is. Tabel 3.1.2 laat zien dat in meting 2 en 3 een op de drie à vier deelnemers die werk heeft gevonden een vast contract heeft gekregen (28%). Ongeveer twee op de drie deelnemers die werk heeft gevonden komt in een tijdelijke baan terecht (67,3%) en slechts 4,5% gaat als zelfstandige aan de slag. Een op de zes deelnemers vindt een uitzendbaan (16%). Veel flexbanen zijn deeltijdse banen dus we vermoeden dat van de deelnemers die een tijdelijke baan vonden dat voor velen een deeltijdbaan was. In de analyses op de CBS data ten behoeve van de definitieve rapportage zal op basis van de inkomensinformatie in de CBS bestanden worden onderzocht in hoeverre er verschillen zijn in

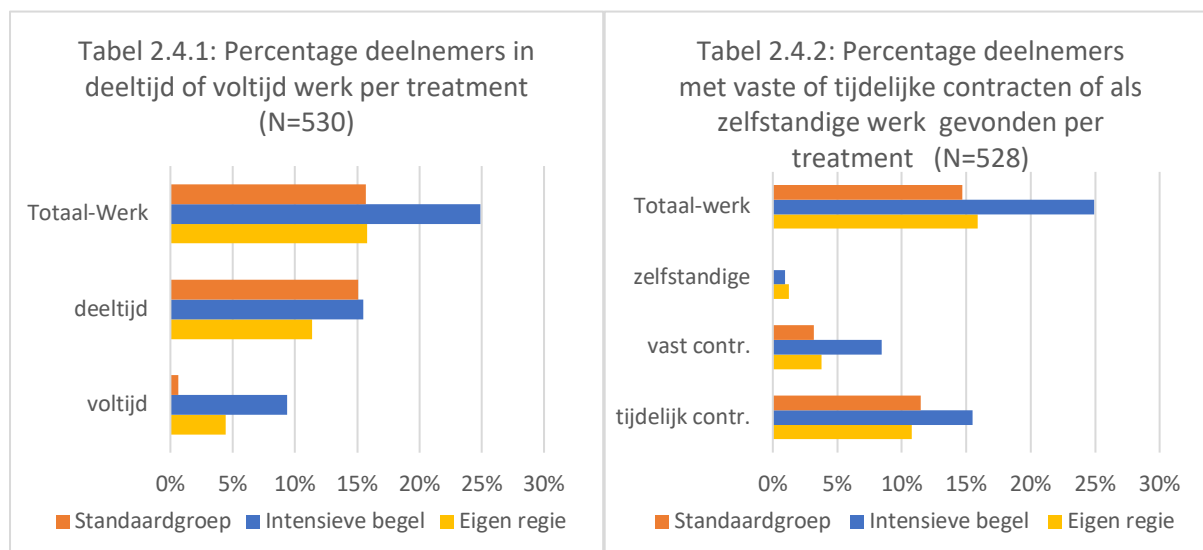
deeltijduitstroom en uitstroom naar flex tussen de twee treatmentgroepen en de standaardgroep en referentiegroep. Dat is met de beschikbare BUS-data vooralsnog niet mogelijk.

Tabel 3.1.2: Type contract van deelnemers die werk vonden sinds de vorige keer, in meting 2 en 3, Apeldoorn/Epe

Type contract	Meting 2+3		
	N=	%	In % deelnemers
Vast contract (onbepaalde tijd)	31	28,2%	5,5%
Tijdelijke werk (contract voor bepaalde tijd)	48	43,6%	8,5%
Oproepkracht	8	7,3%	1,4%
Uitzendkracht	18	16,4%	3,2%
Zelfstandige/vrij beroep	5	4,5%	1,0%
Aantal deelnemers meting 2+3=566	110	100%	19,6%

Bron: Vragenlijst eerste vervolgmeting deelnemers , TiU, 2019

Op basis van deze gegevens is het aandeel berekend van de verschillende type banen (voltijd, deeltijd en vast versus tijdelijk) in de drie treatment groepen in Apeldoorn/Epe (zie Fig. 2.1 en 2.2). Het aandeel deeltijdbanen in de eigen regie groep is wat kleiner is dan in de controlegroep en de intensieve groep. De standaard -of controlegroep en intensieve begeleiding hebben relatief de meeste tijdelijke banen maar de intensieve groep ook de meeste voltijdbanen in vergelijking met de controlegroep (verschil is significant).



Bron: Vragenlijst eerste en tweede vervolgmeting deelnemers Apeldoorn/Epe, TiU, 2020

Fig. 2.1-2.2.: Aandeel type banen per treatmentgroep in het Vertrouwensexperiment Apeldoorn/Epe, 2017-2019

Bij de intensieve begeleidingsgroep gaan in vergelijking met de controlegroep ook de meeste mensen aan het werk. Hierbij zij aangetekend dat omdat het geen volledig beeld geeft we voorzichtig moeten zijn met het trekken van definitieve conclusies. Het beeld dat hieruit ontstaat suggereert nochtans dat de intensieve groep in Apeldoorn/Epe succesvol is geweest in het aan het werk krijgen van deelnemers, zowel in tijdelijke en deeltijdbanen als in voltijdbanen. Het grootste deel van de banen zijn zoals eerder geconstateerd tijdelijke banen en niet bekend is hoe duurzaam deze banen zijn. Daarvoor is de tijdspanne van het onderzoek te kort geweest.

3.1.1 Resultaten: volledige uitstroom naar werk per treatmentgroep

Op basis van de BUS-gegevens is een analyse verricht op voltijdse (minimaal 27 uur) uitstroom naar werk en daarmee uitstroom uit de bijstand. Dit is gebeurd op basis van de registratie in de BUS van de reden van beëindiging van de uitkering. In hoeverre deze registratie een betrouwbare bron is om volledige uitstroom naar werk te achterhalen, is niet onderzocht. We nemen voorsnag aan dat de administratie van de gemeente betrouwbare informatie verschaft over uitstroom uit de uitkering. Mensen waarvan de uitkering is beëindigd vanwege ziekte of arbeidsongeschiktheid zijn evenwel ook opgenomen in deze variabele waardoor een volledig zuivere meting van de uitstroom naar betaald werk niet goed mogelijk was en de uitstroompercentages licht zijn overschat. Vergelijking van de BUS-cijfers met de administratie van volledige uitstroom die door het Vertrouwensexperiment in Apeldoorn/Epe zelf is bijgehouden laat evenwel zien dat er slechts kleine verschillen zijn tussen beide registraties in de uitstroom naar betaald werk¹⁰.

Peildatum en mutaties in de BUS-registratie

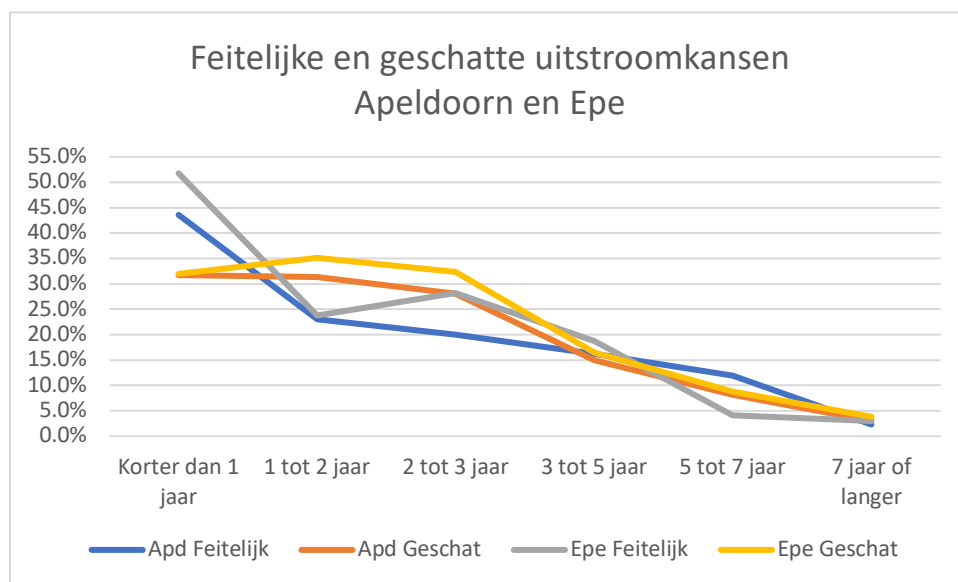
De peildatum van uitstroom is 1 oktober 2019. Alle beëindigingen van de bijstandsuitkering volgens de Participatiewet vanwege uitstroom naar ziekte arbeidsongeschiktheid of werk tot 1 oktober 2019 zijn in de analyses verwerkt. Dat betekent dat de observatie van de uitstroom naar werk een periode van twee jaar omvat. Hiervoor is gebruik gemaakt van de BUS-informatie ultimo oktober 2019 en in het bijzonder de datum van beëindiging van de uitkering welke op dat moment in de administratie is verwerkt. Gemeenten passen evenwel mutaties toe op de administratie tot 3 maanden na de eerste registratie. Dit heeft te maken met het herzien van het recht op uitkering, overgang naar andere typen bijstandsregelingen (bijvoorbeeld van alleenstaande naar samenwoning), en verandering in onder andere de persoonlijke inkomens -en vermogenssituatie van de bijstandsgerechtigde. De hier gebruikte data zijn dus voorlopige data van de gemeente Apeldoorn/Epe.

Volledige uitstroom naar werk

In de Tabellen 3.1.3 en 3.1.4 presenteren we de cijfers over de bereikte uitstroompercentages naar werk voor elk van de treatmentgroepen, de standaardgroep en de referentiegroep. Hierbij is nog niet gecorrigeerd voor compositieverschillen, ofwel verschillen in samenstelling tussen de treatment groepen.

¹⁰ De cijfers van de gemeente zelf registreren een uitstroom van 99 voor de deelnemers exclusief afmeldingen terwijl wij op basis van de BUS-data per 1 oktober 2019 een uitstroom naar werk registreren van 96 voor alle deelnemers exclusief afmeldingen en 106 inclusief afmeldingen.

Eerder zagen we al dat kortere duren licht zijn oververtegenwoordigd en langere duren licht zijn ondervertegenwoordigd in Apeldoorn en Epe. Dit vertekent enigszins de uitkomsten aangezien de kans op uitstroom sterk samenhangt met de duur van de uitkering. Daardoor zullen de uitstroompercentages iets hoger zijn dan wanneer er geen vertekening naar duur was opgetreden. Uit de regressieanalyse die we later presenteren zal blijken dat deze vertekening niet erg groot is. In Figuur 1 geven we de uitstroomkans weer voor de doelpopulatie dat zijn dus alle mensen die voldoen aan de toelatingscriteria, en op 1 oktober 2017 een lopende Participatiewet uitkering hadden of tussen 1 oktober 2017 en 1 juli 2018 in de PW uitkering instroomden. Hiertoe behoren dus ook de 492 deelnemers aan het experiment (412 Apeldoorn en 81 Epe) inclusief de 88 (79 Apeldoorn en 9 Epe) afmeldingen. De afmeldingen zijn meegenomen in de analyse omdat zij zich hebben aangemeld als deelnemer ook al hebben zij de treatment niet of slechts zeer gedeeltelijk ondergaan. Dit wordt een 'intention-to-treat' analyse genoemd. Iedereen die zich heeft aangemeld en tot de doelpopulatie behoort en daarna aselekt is toegewezen (gerandomiseerd) aan een treatment groep wordt in principe in deze analyse gevolgd. Naast deze ITT-analyse voeren we een analyse uit op de deelnemers die de treatment daadwerkelijk hebben ontvangen. Figuur 1 laat een sterke en bijna lineaire daling van de uitstroomkans zien naar werk bij volledige uitstroom uit de uitkering (27 uur bij het minimumloon) in zowel Apeldoorn als Epe hoewel in Epe met wat meer knikken in de curve. De gemiddelde duur van de PW uitkering waarop Figuur 1 gebaseerd is bedraagt ongeveer 5 jaar in Apeldoorn en 4,5 jaar in Epe.



Figuur 3.1.1. Uitstroomkansen naar voltijds werk naar duur uitkering 2017-2019, Apeldoorn (n=3822) en Epe (n=473)

De gemiddelde uitstroomkans over deze tweejaarlijkse periode is 19% in Apeldoorn en 21.6% in Epe. De uitstroomkans is in Epe significant hoger dan in Apeldoorn. Eerder berekenden we voor de jaren 2016-2018 een gemiddelde uitstroomkans naar werk in Apeldoorn/Epe van 15% hetgeen betekent dat in 2018 en 2019 de uitstroomkans is gestegen van 15% naar bijna 20%. Dit geeft de verbeterde

arbeidsmarktsituatie weer sinds die tijd. De geschatte of voor een aantal achtergrondkenmerken gecorrigeerde uitstroomkans vlakkt het verloop van de curve met name bij de kortere duren af maar vermindert ook de verschillen tussen Epe en Apeldoorn. De geschatte kansen wijken voor het overige nauwelijks af van de feitelijke. Dit komt door de randomisatie van de drie treatmentgroepen welke kennelijk goed is gelukt.

In Tabel 3.1.3 en 3.1.4 worden de ongecorrigeerde uitstroompercentages naar voltijds werk per treatmentgroep gegeven. Daarbij geeft tabel 3.1.3 de informatie inclusief en exclusief de 88 afmeldingen voor Apeldoorn en Epe. In tabel 3.1.4 maken we een verder onderscheid tussen deelnemers in het zittende bestand en nieuwe instroom deelnemers.

De resultaten gepresenteerd in Tabel 3.1.3 laten bij de ITT-analyse (inclusief afmeldingen) een positief verschil zien tussen de uitstroompercentages van eigen regie vergeleken met de standaardgroep in Apeldoorn (+4,8%). In Epe is er een veel groter verschil tussen beide groepen maar negatief (-7,2%). De standaardgroep realiseert in Epe een hogere uitstroom dan de andere treatments. Hierbij moet worden opgemerkt dat de consulent die de standaardgroep in Epe onder zijn hoede heeft ook een van beide treatmentgroepen in Apeldoorn heeft. In hoeverre dit van invloed is geweest valt moeilijk te zeggen. De intensieve begeleidingsgroep realiseert de hoogste uitstroom in Apeldoorn (+7,4%) maar een lagere uitstroom in Epe (-3,8%) in vergelijking met de standaardgroep. In de analyse exclusief afmeldingen zien we in beide gemeenten de verschillen met de standaardgroep toenemen; het verschil is sterker positief in Apeldoorn (+11,3%) en iets sterker negatief in Epe (-4,1%). Vergelijken we in Apeldoorn de twee treatments, inclusief de afmeldingen, nu met de referentiegroep in plaats van de standaardgroep (de drie rijen links onder) dan valt op dat de beide alternatieve treatments een 11,8% hogere uitstroom naar werk realiseren dan de referentiegroep. Exclusief afmeldingen wordt het verschil nog groter en neemt het toe naar 13,6%. Dit is een fors verschil.

Tabel 3.1.3. Voltijdse uitstroom naar werk (vtwerk) tussen 1 oktober 2017 en 1 oktober 2019 inclusief en exclusief afmeldingen, aantallen en in %, Apeldoorn en Epe

Totaal (zittend bestand + nieuwe instroom)	Apd-Inclusief afmeldingen		Epe-Inclusief afmeldingen		Apd-Exclusief afmeldingen		Epe-Exclusief afmeldingen	
Treatment	N=	VTWerk	N=	VTWerk	N=	VTWerk	N=	VTWerk
Eigen regie	158	28.5%	26	15.4%	128	29.7%	22	13.6%
Intensieve begeleiding	167	31.1%	32	18.8%	141	33.3%	30	20.0%
Standaardgroep	156	23.7%	31	22.6%	141	22.0%	29	24.1%
Referentie groep	4125	18.0%	504	22.2%	4125	18.0%	504	22.2%
Treatments	325	29.8%	58	17.2%	269	31.6%	52	17.3%
Deelnemers	481	26.8%	89	19.1%	410	28.3%	81	19.8%

Bron: BUS-gegevens Vertrouwensexperiment Apeldoorn/Epe, situatie per 1 oktober 2019

In Epe zien we het omgekeerde effect. Inclusief afmeldingen is het verschil tussen beide treatments en de referentiegroep 5% in de min; hetzelfde percentage geldt exclusief de afmeldingen. Het is vooralsnog onduidelijk waarom de resultaten voor Epe, voor de eigen regie en intensieve begeleidingsgroep, zo ongunstig afwijken van de resultaten voor Apeldoorn.

De standaardgroep realiseert een hogere uitstroom naar werk in Epe dan elke andere groep. Het verschil met de referentiegroep is echter niet zo groot. Van belang is op te merken dat de uitvoering van het experiment in Epe voor een belangrijk deel in handen was van consultants in dienst van de gemeente Apeldoorn. Daar vonden we wel een fors verschil tussen standaardgroep en referentiegroep. Het zou kunnen zijn dat daarnaast onbekendheid met de lokale context een rol heeft gespeeld in de slechte uitkomsten voor de intensieve groep. Wellicht is daardoor de referentiegroep de beste vergelijkingsgroep voor de beide treatments aangezien de reguliere treatment in de referentiegroep niet is beïnvloed door de dialoog, afstemming en samenwerking van de consultants betrokken bij het experiment.

In Tabel 3.1.4 bespreken we de resultaten voor het zittend bestand en de nieuwe instroom apart maar niet meer per gemeente. In Epe zijn de aantallen daarvoor te klein om betrouwbare gegevens te kunnen presenteren.

Tabel 3.1.4. Volledige uitstroom naar werk tussen 1 oktober 2017 en 1 oktober 2019 inclusief en exclusief afmeldingen, Apeldoorn/Epe naar zittend bestand versus nieuwe instroom per 1 oktober 2017

Groepen	Inclusief afmeldingen			Exclusief afmeldingen		
Treatments incl. afm.	Zittend bestand	Nieuwe instroom	Totaal	Zittend bestand	Nieuwe instroom	Totaal
Eigen regie	21.9%	40.9%	26.1%	21.9%	44.4%	27.3%
Intensieve begeleiding	26.7%	34.0%	29.0%	29.1%	36.4%	30.4%
Standaard groep	21.6%	30.8%	23.5%	20.3%	29.7%	22.3%
Referentie groep	16.5%	25.1%	18.4%	16.5%	25.1%	18.4%
Totaal (N=)	17.3% (4039)	26.2% (1166)	19.3% (5205)	17.2% (3982)	26.3% (1150)	19.2% (5132)
Treatments	24.3% (283)	37.6% (101)	27.9% (384)	25.7% (241)	40.4% (79)	29.4% (320)
Deelnemers totaal (N=)	23.4% (431)	35.3% (139)	26.1% (570)	23.8% (374)	38.1% (118)	27.2% (492)

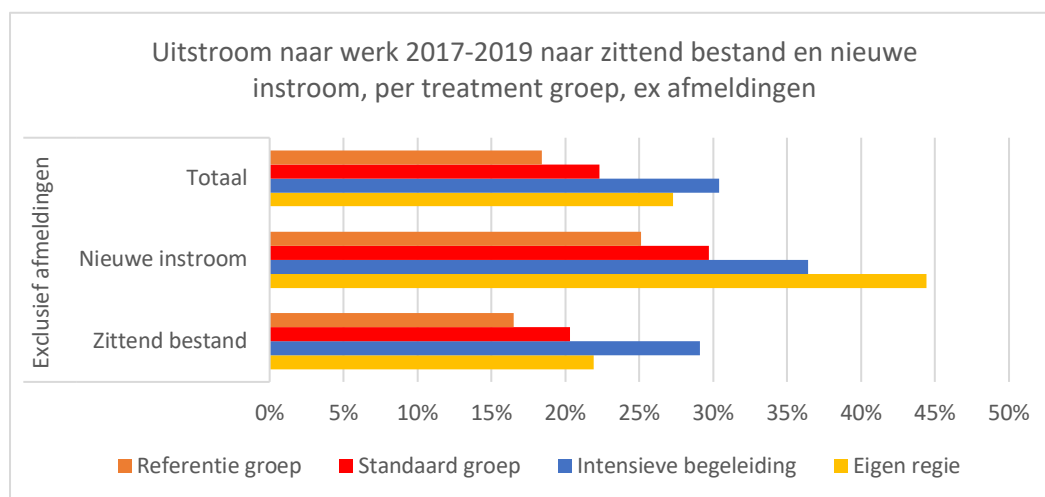
Bron: BUS-gegevens Vertrouwensexperiment Apeldoorn/Epe, situatie per 1 oktober 2019

Uit tabel 3.1.4 leiden we het volgende af:

- De uitstroomkansen voor nieuwe instromers (instromers tussen 1 oktober 2017 en 1 juli 2018) zijn aanmerkelijk groter (26%) zijn dan voor mensen in het zittende bestand (17%). Voor de deelnemers loopt dat verschil tussen zittend bestand en nieuwe instroom verder op naar 12% (23% versus 35%).

- De verschillen tussen de twee treatments en de standaardgroep zijn groter voor nieuwe instroom (6,8%) dan voor het zittende bestand (2,4%).
- In beide gemeenten tezamen vinden we de grootste effecten bij eigen regie voor nieuwe instroom en bij intensieve begeleiding voor het zittende bestand vergeleken met de standaard -en referentiegroep. Bij eigen regie is het verschil voor nieuwe instroom 10,1% vergeleken met de standaardgroep en 15,8% vergeleken met de referentiegroep. Bij intensieve begeleiding is het verschil voor het zittende bestand 5,1% met de standaardgroep en 10,2% met de referentiegroep. Deze verschillen zijn significant en groter dan verwacht.
- Het verschil tussen voltijdse uitstroom naar werk in eigen regie en referentiegroep is 10-15% voor nieuwe instromers (ingestroomd tussen oktober 2017 en oktober 2018). Deze nieuwe instromers komen in de huidige arbeidsmarktsituatie relatief ten opzichte van het zittende bestand gemakkelijk aan het werk.

In Figuur 3.1.2 geven we de resultaten weer voor de verschillende groepen uit tabel 3.1.4 waarin de afmeldingen niet zijn meegenomen. De grafiek totaal geeft het gemiddelde per treatment weer over de beide strata. De uitstroomkansen voor nieuwe instromers zijn voor alle groepen hoger dan voor mensen in het zittende bestand. Vergelijken we de treatments dan zien we dat de eigen regie groep een hogere uitstroom heeft gerealiseerd bij de nieuwe instroom en intensieve begeleiding bij het zittende bestand. Dit is overeenkomstig de verwachting dat intensieve begeleiding beter tot zijn recht zou komen bij het zittende bestand en dat eigen regie beter tot zijn recht zou komen bij de nieuwe instromers met betere perspectieven op de arbeidsmarkt.



Bron: Gegevens BUS en Vertrouwensexperiment Apeldoorn/Epe; bewerking door TiU/Tranzo

Figuur 3.1.2. Uitstroom naar betaald werk voor zittend bestand en nieuwe instroom naar treatmentgroep en referentiegroep, Apeldoorn/Epe, 2017-2019

Verder valt op dat de standaardgroep met de reguliere aanpak wat uitstroom naar werk betreft flink achterblijft bij de intensieve begeleidingsgroep maar ook bij de eigen regiegroep maar dan vooral bij nieuwe instromers. Wellicht heeft hier enige zelfselectie plaatsgevonden bij de aanmeldingen en afmeldingen. Alle groepen inclusief de standaardgroep realiseren een hogere uitstroom dan het geval is

in de referentiegroep die evenals de standaardgroep de reguliere treatment heeft gekregen. De standaard treatment is kennelijk niet goed vergelijkbaar geweest met de standaard treatment in de referentiegroep. Dit zou kunnen wijzen op het zogenaamde “Hawthorne” effect dat de standaard treatment onder invloed van het experiment en het onderzoek meer is gaan lijken op de aanpak in de andere treatments. Zoals gezegd zijn de uitstroomcijfers mogelijk een beetje overschat door een lichte oververtegenwoordiging van korte duren en een lichte ondervertegenwoordiging van langere duren. Hieronder willen we graag corrigeren. Hieronder bespreken we daarom de methode en de uitkomsten daarvan die is gebruikt om daarvoor te corrigeren.

Modelmatige schatting van de uitstroomkans

Om te corrigeren voor compositieverschillen tussen de groepen ook wat duur betreft hebben de onderzoekers bij de analyse van voltijdse uitstroom naar betaald werk gebruik gemaakt van een analysemodel dat zoveel mogelijk vergelijkbaar is gemaakt met het model dat in het landelijk overleg van onderzoekers is overeengekomen. In Tabel 3.15 worden de gebruikte variabelen weergegeven.

Tabel 3.1.5. Aantal, gemiddelde en standaardafwijking gebruikte variabelen in het analysemodel volledige uitstroom naar werk, Apeldoorn/Epe, 2017-2019

Naam model variabele	Omschrijving	Frequentie (N=)	Gemiddelde	Standaardafwijking
uitstrwerk	Reden van uitstroom (werk, ziekte)	5212	0.19	0.38
ovexgroep4	Onderzoeksgroep (treatments, standaardgroep, referentiegroep)	5218	3.78	0.68
duuruitk_t0	Duur van uitkering PW bij start experiment ¹⁾	5198	3.10	4.95
nkeeruitk_t0	Aantal keer uitkering voor start experiment	5193	1.17	0.46
lft_t0	Leeftijd bij start	5190	43	14
geslacht_t0	Geslacht (vrouw=1; man=0)	5212	0.52	0.5
buitnat_t0	Nationaliteit (Nederlands of buitenlands)	5218	0.19	0.39
leefv_t0	Leefvorm (alleenstaande, alleenstaande ouder, samenwonend)	5212	2.12	0.62
huisv_t0	Huisvesting (eigenaar, huurder, inwonend, overig)	5212	2.2	0.63
nizb_t0	Zittend bestand of nieuwe instroom ¹⁾	5211	0.78	0.42
gemeente	Dummy voor gemeente Epe t.o.v. Apeldoorn (=0)	5205	0.11	0.32

Noot: ¹⁾ De meerderheid van de mensen in de treatmentgroepen komt uit het zittende bestand (dus heeft een lopende uitkering per 1 oktober 2017). In het model is de log van de duur genomen aangezien de relatie tussen duur en uitstroom naar werk log lineair is. De variabele nkeeruitkt0 was niet significant en is weggelaten uit het model.

Bron: BUS-data gemeente Apeldoorn/Epe, 2017-2019

Het LOEP-model is afgestemd op de beschikbare informatie in de CBS-data, niet alleen de BUS-data waarvan het CBS ook gebruik maakt, maar ook andere registratiedata die het CBS gebruikt. In het model

dat wij gebruiken voor Apeldoorn/Epe zijn zoveel mogelijk dezelfde variabelen gebruikt als in het LOEP-model met dezelfde categorisering. Bij bepaalde variabelen bleek dat moeilijk te realiseren zoals bij opleiding door een andere indeling in de BUS data. Hierna presenteren wij de uitkomsten van het model. De afhankelijke variabele is voltijdse uitstroom naar betaald werk. De onafhankelijke variabelen zijn maand geen informatie in de BUS-data beschikbaar was. Dan is de dichtstbijzijnde maand genomen waarvoor wel informatie beschikbaar was. De gemiddelde duur bedraagt bij aanvang van het experiment meer dan 4 jaar (6 jaar minus 2 jaar duur experiment).

Resultaten

Tabel 3.1.6 geeft de resultaten van de logistische regressie weer. Het model verklaart 18% van de variantie (pseudo R^2) en is significant. Het model corrigeert zoals gezegd voor duurverschillen tussen de groepen maar ook voor andere compositie verschillen die te maken hebben met geslacht, leeftijd, nationaliteit, leefvorm, huisvesting en herinstroom.

Tabel 3.1.6. Resultaten logistische regressie: geschatte uitstroomkansen naar treatment groep, inclusief en exclusief afmeldingen, plus treatment-effecten in procentpunten

	Apeldoorn				Epe			
Totaal	Inclusief afmeldingen		Exclusief afmeldingen		Inclusief afmeldingen		Exclusief afmeldingen	
Treatment	N=	P-werk	N=	P-werk	N=	P-werk	N=	P-werk
Eigen regie	151	29.1%	125	28.6%	23	17.4%	20	16.0%
Intensieve begeleiding	156	30.8%	134	30.8%	32	18.8%	30	19.1%
Standaardgroep	151	24.5%	136	24.3%	30	23.3%	28	24.5%
Referentie groep	4100	18.0%	4100	18.0%	496	22.0%	496	22.0%
Treatments	307	30.0%	259	29.7%	55	18.1%	50	17.9%
Deelnemers	458	28.2%	395	27.9%	8ii5	20.0%	78	20.3%
Totaal	4558	19.0%	298	32.3%	581	21.7%	4495	21.7%
Pseudo R^2 = Wald $\chi^2(22)$ Prob > χ^2 N=	0.18 639.6 0.000 5139	<i>Contrast-analyse</i> ER--SG (ni)* IB--SG (zb)* ER--RG** IB--RG (zb) *** SG--RG (zb)*** ER+IB--RG**						

Noot: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$; Logit regressie model met controlevariabelen leeftijd, leeftijd kwadraat, nationaliteit, geslacht, log duur uitkering bij start, nkeer uitkering bij start, leefvorm bij start, huisvesting bij start. ER=eigen regie; IB=Intensieve begeleiding; SG=Standaard/controlegroep; zb=zittend bestand; ni=nieuwe instroom.
Bron: Bewerking BUS data, Tilburg University, 2019.

Vanwege de niet-lineariteit in de relatie tussen uitstroom naar werk en leeftijd (omgekeerde U-vorm) en duur uitkering (log-lineair) hebben we tevens leeftijd in het kwadraat toegevoegd en hebben we de natuurlijke logaritme van duur gebruikt. We hebben tevens een model geschat voor de vergelijking tussen de drie groepen en dus niet met de referentiegroep en dan blijkt dat dit andere model significant is maar geen significante of betekenisvolle verschillen laat zien tussen de treatment groepen en de standaardgroep op deze uitkomstmaat. In Tabel 3.1.6 hebben we op basis van het model de voorspelde uitstroomkansen geschat voor Epe en Apeldoorn apart. De verschillen met de ongecorrigeerde cijfers zijn niet zo groot. De eigen regie groep doet het iets beter in Epe dan op basis van de ruwe cijfers maar nog steeds minder goed dan de standaardgroep of de referentiegroep.

Onderaan in de tabel geven we de technische resultaten van de logit regressies weer. We hebben twee regressies gedraaid een met de standaardgroep als vergelijking en de andere met de referentiegroep als vergelijking. De grootte van de effecten weergegeven in procentpunten ten opzichte van de standaardgroep en de referentiegroep verandert niet in deze multivariate analyse vergeleken met de ruwe scores. Bij de analyse inclusief afmeldingen (ITT-analyse) zijn de effecten ten opzichte van de standaardgroep nog steeds zo'n 5 procentpunten tegenover 10-12 procentpunten in vergelijking met de referentiegroep.

Contrasten

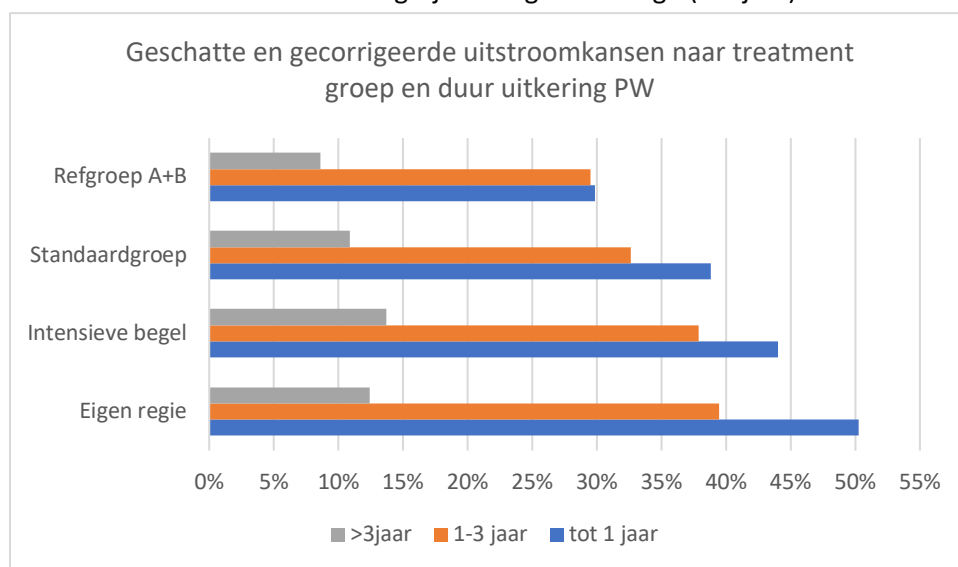
Voor de contrastanalyse hebben we twee modellen geschat, een met treatment en de 'nieuwe instroom'- 'zittend bestand' (nizb) variabele als hoofdeffect en een met interactie-effecten tussen de treatment groep en deze nizb-variabele. De contrastanalyse onderaan de tabel geeft weer welke verschillen tussen de groepen significant zijn op welk significantieniveau (betrouwbaarheidsniveau). We vinden significante effecten voor zowel eigen regie als intensieve begeleiding ten opzichte van de standaardgroep. We vinden opvallend genoeg ook een sterk effect van de standaard treatment ten opzichte van de referentiegroep maar alleen voor het zittende bestand. Dit is een sterke indicatie dat de standaard treatment is veranderd gedurende het experiment. De sterkste effecten treden nochtans op bij intensieve begeleiding van het zittende bestand in vergelijking met de referentiegroep. Betekenisvolle effecten bestaan er ook bij de eigen regie groep zowel voor het zittende bestand als voor nieuwe instroom maar ten opzichte van de referentiegroep. Ten opzichte van de standaardgroep is alleen de uitstroom voor nieuwe instroom betekenisvol hoger. Tot slot zien we dat de twee treatments gezamenlijk significante positieve effecten (op 95% betrouwbaarheidsniveau) laat zien ten opzichte van de referentiegroep maar niet ten opzichte van de standaardgroep.

3.1.2 De invloed van duur nader onderzocht: matching modellen

De resultaten laten forse verschillen zien tussen de drie treatment groepen en de referentiegroep. Vanwege zelfselectie bij de aanmeldingen en bij de afmeldingen kunnen er verschillen ontstaan tussen de doelpopulatie en de twee referentiegroepen voor zittend bestand en nieuwe instroom. Uit de analyse van deze verschillen bij het onderzoek naar de representativiteit van de doelpopulatie constateerden we dat deze verschillen erg meevallen, met name de verschillen tussen de drie treatment groepen maar ook de verschillen met de referentiegroep. We vonden een lichte ondervertegenwoordiging van lange duren en een lichte oververtegenwoordiging van korte duren. Duren van 1-3 jaar bleken evenredig te zijn

vertegenwoordigd. Tevens zijn er relatief weinig mensen van buitenlandse nationaliteit in de populatie. De vraag stelt zich nochtans of het model in staat is om op correcte wijze te corrigeren voor de verschillen tussen de treatment groepen aan de ene kant en de referentiegroep. In de logit regressiemodellen vonden we een sterk effect van duur in de bijstand; met elk extra jaar van verblijf in de bijstand neemt de kans op uitstroom naar voltijds werk af met 1,4%. Het effect was een van de sterkste effecten, negatief dus en significant op 95% niveau. De eerder weergegeven uitstroomkansen hebben daarvoor zo goed mogelijk gecorrigeerd.

In Figuur 3.1.3. presenteren we de geschatte en voor duurverschillen en verschillen in andere achtergrondkenmerken gecorrigeerde uitstroomkansen naar duur en treatmentgroep. Na correctie voor duur blijven de gevonden verschillen in uitstroomkansen met de referentiegroep overeind voor alle duren dus ook voor duren die evenredig zijn vertegenwoordigd (1-3 jaar).



Bron: Bewerking BUS data, Tilburg University, 2019.

Fig. 3.1.3: Geschatte uitstroomkansen naar voltijds werk naar duur en treatment groep, Apeldoorn en Epe, (n=5118)

Rol van klantmanager

Kennelijk speelt de duur minder een rol dan andere factoren die het verschil kunnen verklaren. Een van die factoren is wellicht de klantmanager zelf. Eenzelfde logit regressie model maar nu alleen geschat op de mensen in de drie treatment groepen en met de standaardgroep als vergelijkingsgroep maar met klantmanager-nummer als verklarende variabele verklaart de verschillen in uitstroomkansen naar werk tussen de groepen ruim de helft beter als hetzelfde model zonder klantmanager-nummer (21% in plaats van 15%). Van belang is mogelijk dat de klantmanagers in het experiment bewust hebben gekozen voor deelname als consultant aan het experiment. Ze zijn om praktische redenen - hoewel het RCT design stricto sensu dat veronderstelt - niet aselekt geselecteerd waardoor de motivatie voor het experiment een rol kan hebben gespeeld. Bovendien zijn de klantmanagers veelal 'dedicated' geweest aan een bepaalde

treatment groep. Dat gold zoals we eerder vermeldden niet voor alle klantmanagers. Die extra motivatie kan wellicht deels ook de effecten verklaren.

Matching: correctie voor compositieverschillen

Ten tweede is er een 'matching' methode gebruikt om de verschillen tussen de treatmentgroepen en de referentiegroepen op het spoor te komen. Indien de populatie van deelnemers door (zelf)selectie niet volledig representatief is op bepaalde aspecten, bijvoorbeeld duur, leeftijd, en buitenlandse nationaliteit, dan kan in de data gezocht worden naar andere mensen die zoveel mogelijk lijken op de deelnemers maar die de treatment niet hebben gehad. We hebben deze methode gebruikt om de resultaten bij de deelnemers te vergelijken met een gematchte groep van mensen die qua duur en andere kenmerken op de deelnemers lijken. Op deze manier kan worden gecorrigeerd voor verschillen in waargenomen kenmerken zoals duur, leeftijd en nationaliteit. We hebben 'een-op-een' matching toegepast, dus voor iedere deelnemer is één andere persoon gezocht in de referentiegroep die op de deelnemer lijkt. Indien meer personen vergelijkbaar zijn wordt degene gekozen die het meest lijkt op de klant. Dat wordt de 'dichtstbijzijnde buur' ('nearest neighbour') genoemd¹¹. Voor de matching zijn zoveel mogelijk dezelfde variabelen gebruikt als in het regressiemodel. In beide methoden (nearest neighbour matching en propensity score matching) werd een betere matching bereikt met het weglaten van de variabele die aangeeft of men tot het zittende bestand of nieuwe instroom behoort. Ook leverde een andere opleidingsvariabele (met vier niveaus) betere 'matching' resultaten op. De vier niveaus zijn laag, midden, hoog en geen opleiding of opleiding onbekend¹². De beide methoden laten toe om het effect van de interventie bij de treated zuiver vast te stellen (ATT-average treatment effect van de treated) maar ook het gemiddelde effect van de treatment voor de gehele populatie van deelnemers en niet-deelnemers. Voor de niet-deelnemers wordt dan op basis van hun kenmerken berekend wat het effect zou zijn geweest als zij ook de treatment hadden ondergaan (ATE-average treatment effect). Tot slot is het mogelijk om het effect te bepalen voor degenen die geen interventie hebben gehad ervan uitgaande dat zij die wel zouden hebben gehad (ATU -average treatment effect van de untreated). In onderzoek wordt veelal de ATT en ATE gebruikt om het effect van een interventie te meten. De ATT en ATE geven dan het verschil in geschatte uitstroomkans weer tussen de beide treatmentgroepen en de controlegroep. De ATE geeft het gemiddelde effect over de deelnemers en niet-deelnemers en is daarom wellicht de beste maatstaf. We vonden zowel betekenisvolle verschillen van de twee treatmentgroepen met de standaardgroep als met de referentiegroep hoewel de laatste verschillen aanmerkelijk groter waren¹³. We hebben de 'nn-matching' en 'ps-matching' modellen geschat. Ten eerste zien we dat de effecten wat betreft de vergelijking van de treatments met de standaardgroep die op basis van het logistisch regressiemodel wel significant waren, dat nu niet meer zijn. De significante verschillen in vergelijking met

¹¹ De matching (zoeken en bepalen van vergelijkbare mensen) kan op basis van een wegingsfunctie van de verklarende variabelen (bijvoorbeeld middels nearest neighbour matching) of op basis van zogenaamde 'propensity scores' (de kans op treatment bepaald met behulp van een logit of probit model met verklarende variabelen) van de covariaten. Beide methoden zijn toegepast met vergelijkbare resultaten.

¹² Laag staat voor enkele klassen lager onderwijs of lager onderwijs; midden staat voor (voorbereidend) middelbaar algemeen vormend en beroepsonderwijs, hoog staat voor hoger beroeps- en wetenschappelijk onderwijs.

¹³ De modellen zijn ook geschat met de standaardgroep als vergelijkingsgroep maar de resultaten bevestigen de uitkomsten van de logistische regressies; er werden geen significante verschillen gevonden tussen beide treatmentgroepen en de standaardgroep.

de referentiegroep blijven wel overeind. Vergelijken we de twee treatments met de referentiegroep dan bedraagt de ATT 6,1%, en de ATE 7,3% in de wegingsmethode. In de 'propensity score' matching methode bedraagt de ATT 4,9% en de ATE 5,0%. Dat betekent dat er op basis van de ATE een 5% tot 7,5% hogere kans bij de treatmentgroepen is om volledig uit te stromen in vergelijking met de referentiegroep. Het effect is minder dan bij de logit regressie resultaten (10-15%) maar niettemin een substantieel effect. De correctie voor compositieverschillen heeft zoals verwacht ertoe geleid dat het effect nu kleiner is. De resultaten van de analyse worden in Tabel 3.1.7 weergegeven. Voordat we deze presenteren laten we zien in welke mate het model in staat is geweest om te corrigeren voor verschillen in duur. In Fig. 3.1.4 (links nn-matching; rechts ps-matching) staan in beide plots links de feitelijke uitstroomkansen naar werk en rechts de gecorrigeerde uitstroomkansen voor de controlegroep (referentiegroep hier) en de beide treatments dus na matching.

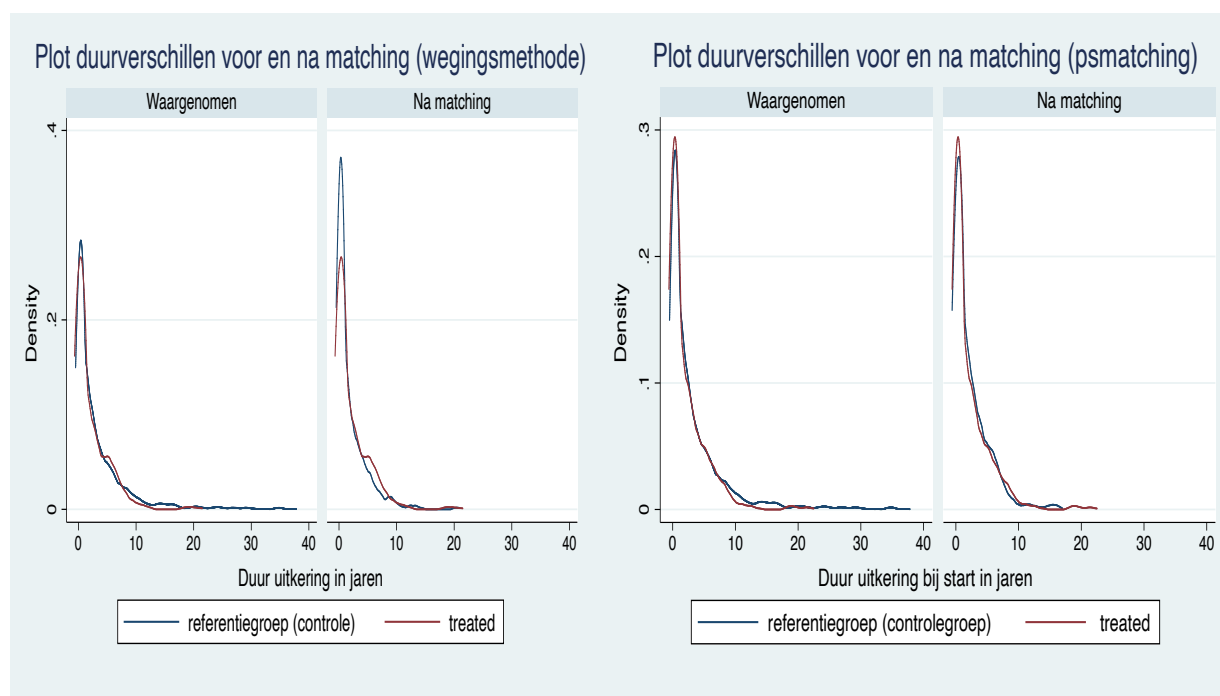


Fig. 3.1.4 Verschillen tussen controlegroep en treatments naar duur uitkering en matchingmethode

De feitelijke kansverdeling in Fig. 3.1.4 laat zien dat er nauwelijks sprake is van duurverschillen tussen de referentiegroep en de deelnemersgroep wat betreft de korte duren. Het verschil zit vooral in de wat langere duren. Na matching zijn de uitstroomkansen van beide treatments naar duur nu nagenoeg hetzelfde (variantie ratio is ongeveer 1) en het verschil in variantie ligt dicht bij 0¹⁴. De matching corrigeert evenwel ook voor andere verschillen in achtergrondkenmerken zoals leeftijd en nationaliteit. Deze zullen ook van invloed zijn op de schatting van de effecten. Nochtans moet worden opgemerkt dat de referentiegroep in tegenstelling tot de standaardgroep niet is gerandomiseerd waardoor we niet zeker weten of de gevonden resultaten causaal geduid kunnen worden. Door de grote aantallen respondenten

¹⁴ In de aanvullende rapportage zullen de technische aspecten worden gepresenteerd.

in de referentiegroep (n=5000) is nochtans de kans op vertekening van de resultaten erg klein. Tabel 3.1.7 geeft de effecten weer voor de twee respectievelijk drie treatment groepen (deelnemers) in vergelijking met de referentiegroep en de standaardgroep. Omdat we eerder geen verschillen vonden tussen de beide treatmentgroepen presenteren we de effecten in de tabel voor beide groepen tezamen.

Tabel 3.1.7. Resultaten matching modellen uitstroomkans werk (voltijds)

	NN-matching ¹⁾		PS-matching		
	ATT	ATE	ATT	ATE	N=
Vershil t.o.v. referentiegroep					
Treatments (eigen regie, intensieve begeleiding)	6,1%**	7,3%**	5,0%*	4,9%**	N=4970
Deelnemers (incl. standaardgroep)	4,6%*	4,1%*	4,2%*	4,9%**	N=5157
Vershil t.o.v. standaardgroep					
Treatments (eigen regie, intensieve begeleiding)	ns	ns	ns	ns	N=564
Model variabelen					
(log) Duur uitkering op t0	x	x	x	x	
Leeftijd op t0	x	x	x	x	
Geslacht (man, vrouw)	x	x	x	x	
Buitenlandse nationaliteit (Nederlands, buitenlands)	x	x	x	x	
Leefvorm (alleenstaand, samenwonend, overig)	x	x	x	x	
Opleidingsniveau (onbekend, laag, midden, hoog)	x	x	x	x	
Wel/geen startkwalificatie	x	x	x	x	
Bias variabelen (duur uitkering, leeftijd en nationaliteit)	x	x			

Noot 1: ns=niet significant op 10% niveau; ¹⁾In het model voor de 'nearest neighbourhood' matching methode is een correctie opgenomen voor bias (vertekening) op duur uitkering, leeftijd en nationaliteit. Beide methoden zijn geschat met Stata 15.1. Voor bepaling van de covariaten in het model is het programma psmatch2 in combinatie met pstest gebruikt.

Noot 2: ***p<0,01; **p<0,05; *p<0,10, getest op basis van een nearest neighbour matching model versus een propensity score matching logit model (N=5157)

Grotere uitstroomeffecten naar werk in Apeldoorn (7-8%)

Op basis van Tabel 3.1.7. concluderen we nu dat er geen significante treatment effecten zijn in vergelijking met de standaard groep. De eerder gevonden significante verschillen vallen weg na matching op basis van een aantal achtergrondkenmerken. De verschillen met de standaardgroep worden kennelijk verklaard door aanwezige compositieverschillen met de standaardgroep. In het logit regressiemodel vonden we nog wel een significant interactie-effect tussen eigen regie en intensieve begeleiding voor het zittend bestand in vergelijking met de standaardgroep. In dat model is ook gecorrigeerd voor compositieverschillen. Het

matching model houdt evenwel geen rekening met zittend bestand of nieuwe instroom waardoor dat effect niet meer zichtbaar is. Op basis van de vergelijking met de standaardgroep moeten we nu concluderen dat de beide treatments weinig effect hebben gehad op de uitstroom naar werk. Maar hoe is dan te verklaren dat er wel een substantieel effect gevonden is in vergelijking met de referentiegroep? Dat valt alleen met elkaar te rijmen indien de standaard treatment zelf is veranderd in deze periode waardoor verschillen in de uitvoering van de drie treatments zijn verminderd. Ten opzichte van de referentiegroep vinden we een sterk positief treatment effect van de beide treatments maar ook van de drie treatments dus inclusief de controle -of standaardgroep. De beide matching methoden verschillen voorts niet veel in uitkomsten. Voor de beide treatment groepen is het effect 5-7% en voor alle drie groepen tezamen 4-5%. Dat is echter een gemiddelde over beide gemeenten.

De effecten op basis van de ruwe scores lieten eerder positieve effecten zien voor Apeldoorn maar negatieve voor Epe. De effecten zijn voor Apeldoorn ook op basis van deze matching methoden groter dan voor Epe. Op basis van ps-matching vinden we voor de twee treatments in Apeldoorn ten opzichte van de referentiegroep effecten ter grootte van 7,8% (ATT) en 6,1% (ATE). Deze zijn dus hoger dan de 5% tot 5,5% voor beide gemeenten tezamen zoals vermeld in Tabel 3.1.7 maar kleiner dan we vonden op basis van de ruwe data. Op basis van matching blijkt dat in vergelijking met de standaardgroep de gemiddelde effecten van de beide treatments voor zittend bestand en nieuwe instroom op de uitstroomkans niet meer significant zijn. In vergelijking met de referentiegroep zijn de effecten op voltijdse uitstroom naar werk echter -ook na correctie op basis van matching- wel significant. Indien we een model schatten met ene interactieterm voor gemeente (Apeldoorn versus Epe) dan is het effect voor Epe niet significant en dus 0.

3.1.3 Conclusie uitstroom naar werk

Samenvattend, laten de resultaten zien dat op basis van de uitkomstmaat van voltijdse uitstroom naar betaald werk (minimaal 27 uur) er op grond van de feitelijke en gecorrigeerde uitstroomkansen (logistische regressie) er betekenisvolle en substantiële verschillen zijn tussen de twee treatmentgroepen en de standaardgroep en tussen de twee treatments met de referentiegroep van niet-deelnemers. Gemiddeld genomen zijn de ruwe of waargenomen uitstroomkansen voor de beide treatmentgroepen 10-15% en voor de drie groepen tezamen 9-12% hoger dan voor de referentiegroep van niet-deelnemers. Dat is een fors verschil. De aangemelde deelnemers zijn aselekt toegewezen aan de verschillende treatments maar door zelfselectie kan de groep van deelnemers qua samenstelling afwijken van de populatie waaruit ze is getrokken. Dat blijkt ook het geval te zijn. Vooral naar duur uitkering, leeftijd en nationaliteit zijn er verschillen tussen de deelnemers en de doelpopulatie ontstaan waarvoor gecorrigeerd moet worden. Daarvoor hebben we matching toegepast waarbij we de twee treatmentgroepen hebben vergeleken met de standaardgroep en met de referentiegroep op basis van dezelfde achtergrondkenmerken als welke zijn gebruikt in het regressiemodel. Uit de matching analyse bleek vervolgens dat de verschillen met de standaardgroep niet meer significant zijn. Vergelijken we de gerealiseerde uitstroomkansen van beide treatment groepen samen of van alle treatments (dus inclusief de standaardgroep) met de referentiegroep (niet-deelnemers) dan zijn de verschillen betekenisvol. Om dat te doen is gebruik gemaakt van twee matching methoden een wegingsmethode en een model

matching methode. Met beide methoden bleek het mogelijk voldoende (100%) geschikte matches te vinden voor de deelnemers onder de niet-deelnemers. Op basis daarvan berekenden we dat er een gemiddeld treatment effect (ATE) van 7 tot 8% bestaat in Apeldoorn en 5-7% voor beide gemeenten tezamen uitgaande van voltijdse uitstroom naar werk van beide treatment groepen in vergelijking met de referentiegroep. Het vermoeden bestaat dat de standaard treatment in het proces van experimenteren mogelijk meer is gaan lijken op de beide andere treatments. De uitstroompercentages zijn in ieder geval aanmerkelijk hoger dan in de referentiegroep (die ook de standaard begeleiding krijgt) en vergelijkbaar met de scores van de beide 'echte' treatment groepen. De wijze van ondersteuning door de klantmanager lijkt van groot belang. Indien we voor de 571 deelnemers waarvoor we betrouwbare informatie hebben (inclusief dus de afmeldingen) een model schatten om de verschillen tussen de groepen te verklaren in uitstroomkansen dan blijkt dat 35% van de verklaarde verschillen (21%) wordt verklaard door de persoon van de klantmanager.

3.1.4 Welbevinden, gezondheid, keuzevrijheid, sociale participatie, zelfredzaamheid, vertrouwen en financiële situatie (armoede)

In de drie vragenlijsten voor de deelnemers (cv-nulmeting, cv-vervolgmeting1 en cv-vervolgmeting 2) zijn vragen gesteld die beogen het experiment te evalueren op de uitkomstmaten die niet direct gerelateerd zijn aan betaald werk zoals welbevinden, subjectieve gezondheid (fysieke en mentale gezondheid), sociale participatie, sociaal en institutioneel vertrouwen, zelfredzaamheid, keuzemogelijkheden, financiële situatie, materiële deprivatie en armoede. In overleg met de andere onderzoekers in LOEP zijn uit een mogelijke lijst van indicatoren waarover informatie in de vragenlijsten beschikbaar was, er negen tot 10 geselecteerd die als leidraad zullen dienen bij de rapportage van de AMvB experimenten. Dat zijn welbevinden, subjectieve gezondheid, sociale participatie, ervaren keuzemogelijkheden (capabilities), zoekintensiteit, zelfredzaamheid, vertrouwen in de ander en in instituties, financiële stress en armoede en/of deprivatie. Deze zullen ook in dit rapport worden gepresenteerd. In de beide tussenrapportages voor Apeldoorn/Epe is over een aantal indicatoren reeds verslag gedaan maar is nog geen informatie gegeven over de vergelijking naar de treatmentgroepen. Deze indicatoren passeren achtereenvolgens de revue hier.

Analysemethode

We hebben nu voor elk van deze indicatoren een regressiemodel (OLS) geschat waarmee we de kans op de uitkomst bepalen waarbij we controleren voor compositieverschillen mogelijk veroorzaakt door verschillen in duur uitkering bij de start, leeftijd, geslacht en nationaliteit tussen de treatments. Daarvoor gebruiken we de BUS informatie over leeftijd, geslacht, opleiding, nationaliteit, leefvorm, huisvesting, duur uitkering bij de start en, aantal keer in uitkering en of men behoort tot het zittende bestand op 1 oktober 2017 of tot de nieuwe instroom die na 1 oktober 2017 is ingestroomd. We hebben de BUS opleidingsinformatie gebruikt over hoogst behaalde opleiding. De modelmatig gecorrigeerde of geschatte cijfers zullen we hierna presenteren en dus niet de feitelijk waargenomen cijfers. Het effect van de correctie is overigens erg klein omdat door de randomisering er kleine verschillen zijn tussen de drie groepen. In de rapportage van de analyse, wordt gebruik gemaakt van een beschrijvende tabel waarin de gemiddelde gecorrigeerde scores voor de treatment groepen op een uitkomstvariabele worden getoond

per meting en in totaal. Deze gemiddelden zijn geschat op basis van de OLS regressieanalyse. Een nadere uitleg over hoe de tabellen gelezen moeten worden staat in paragraaf 3.2.3. Of een effect betekenisvol is of niet en met welke mate van betrouwbaarheid wordt aangegeven met sterretjes; drie sterretjes staat voor 99% betrouwbaarheid twee sterretjes voor 95% en een sterretje voor 90% betrouwbaarheid. Er zijn steeds twee modellen geschat een zonder interactie-effect en een met interactie-effecten tussen treatment en meting (zie Bijlage A voor de regressieresultaten). De effecten in de randtotalen geven op basis van het model zonder interacties de effecten weer van verschillen over de tijd (rijtotalen onder) of van verschillen tussen elk van beide treatments en de standaardgroep gemiddeld over de tijd (kolomtotalen rechts). We zijn met name hierin geïnteresseerd (zie ook Tabel A1.1 in Bijlage A). De sterretjes bij de scores in de andere cellen geven op basis van de schattingen in het interactiemodel weer in hoeverre er een betekenisvol effect is van de treatment ten opzichte van de standaardgroep in de nulmeting bovenop het effect dus dat er was aan het begin van het experiment (zie Tabel A1.2 in Bijlage A). Aangezien aan het begin de treatment de meeste effecten nul zijn (zoals ook vastgesteld) geven deze veelal de volledige treatment effecten weer. In het model zijn naast de treatment variabele controlevariabelen opgenomen maar tevens een variabele die de persoon van de klantmanager weergeeft.

Gezondheid en welbevinden

In de vragenlijsten zijn vragen gesteld die te maken hebben met de gezondheid -en welzijnsbeleving van de klant. Het welbevinden van de deelnemer wordt gepeild aan de hand van vragen over de tevredenheid met het leven (schaal van 0 tot 10), de mate waarin men het leven betekenisvol of zinvol acht (schaal van 0 tot 10) en de mate waarin men zich gelukkig voelt (op een schaal van 1 tot 5)¹⁵.

In Fig. 3.1.5 presenteren we de waargenomen (WB-links) en voor compositieverschillen gecorrigeerde scores (Pwb-rechts) op welbevinden (het gemiddelde van tevredenheid, geluk en betekenisvol leven). De waargenomen en gecorrigeerde scores op welbevinden wijken nauwelijks af hetgeen laat zien dat compositieverschillen dankzij de randomisering (aselecte toewijzing aan de groepen) betrekkelijk weinig invloed hebben. In de nulmeting zijn de feitelijke scores op welbevinden gelijk voor de standaardgroep en de eigen regie groep maar iets hoger voor de intensive begeleidingsgroep (0.05 verschil) om vervolgens toe te nemen in de metingen daarna. Een gering verschil tussen de groepen bij de start duidt erop dat er geen zelfselectie op welbevinden heeft plaatsgevonden voor het experiment.

Figuur 3.1.5 doet vermoeden dat er een stijging plaatsvindt bij vooral de groep intensieve begeleiding ten opzichte van de andere twee treatments. De stijging is echter niet significant aangezien alle drie groepen een stijging in welbevinden laten zien. Over de tijd gezien is er sprake van een stijging maar niet specifiek bij een of andere treatment groep. De OLS regressieanalyse (zie tabel 3.1.8.) geeft uitkomst over welke effecten betekenisvol zijn.

¹⁵ Deze drie vragen zijn vervolgens gebruikt om een gemiddelde score te berekenen van welbevinden waarbij de schaal van geluk eerst omgecodeerd werd naar 0 tot 10.

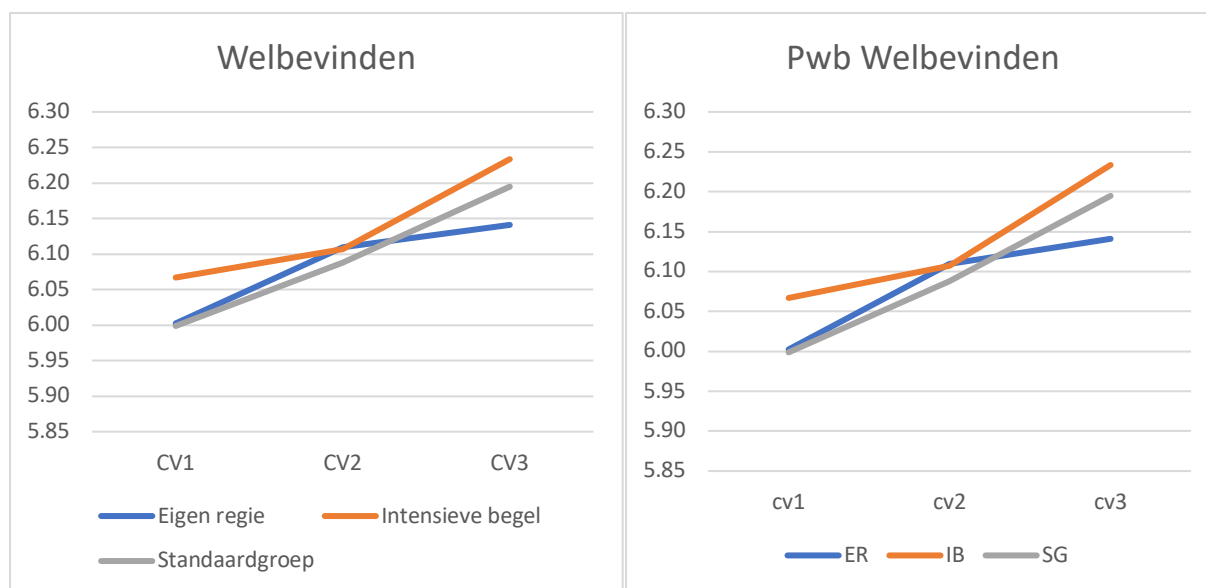


Fig. 3.1.5: Waargenomen en gecorrigeerd niveau van welbevinden (geschatte kans) in de drie metingen
 Noot: WB=Welbevinden; Pwb=Geschatte kans op welbevinden; ER=Eigen Regie; SG=Standaard groep; IB=Intensieve begeleiding

Bron: Cliëntvragenlijsten, nulmeting, meting2 en 3

Ten eerste laat die analyse zien dat over de tijd er bij alle treatments een positieve ontwikkeling is in het welbevinden van de deelnemer maar dat dat niet ligt aan een specifieke treatment. We vinden geen betekenisvolle verschillen tussen de treatments voor de verandering in welbevinden over de tijd (zie de kolomtotalen over alle drie metingen tezamen). Met andere woorden, in alle treatments neemt het welbevinden toe over de tijd gezien, maar niet meer of minder in de ene of de andere treatment. Dus ook niet in vergelijking met de standaardgroep hetgeen wel een van de belangrijkste hypothesen was. Dat is te meer opvallend omdat op grond van de literatuurstudie voorafgaand aan het onderzoek betekenisvolle effecten verwacht werden. Heeft dit te maken met de uitvoering van de treatment of met de treatment zelf? Het welbevinden in de standaardgroep blijkt betekenisvol toe te nemen vanaf meting 2. Dat het welbevinden betekenisvol toeneemt bij de standaardgroep is inderdaad opvallend en doet de vraag rijzen of de treatment hetzelfde is gebleven of is veranderd gedurende de experimentele periode. In het eerste geval zouden de welbevinden scores min of meer hetzelfde zijn gebleven als aan het begin van het experiment tenzij andere factoren zoals de economische situatie of de arbeidsmarktsituatie die in deze 2 jaar zijn verbeterd ook geleid hebben tot een stijging van het niveau van welbevinden. Helaas ontbreken cijfers over welbevinden bij de referentiegroep om dit te kunnen testen. Wel geven de cijfers over welbevinden en geluk in Nederland voor de onderkant van de inkomensverdeling geen reden om een stijging te verwachten. Het SCP constateert zelfs in een recente publicatie dat de subjectieve levenssatisfactie score van mensen met een score van 6 of lager in de afgelopen jaren (hetgeen bij bijstandsgerechtigden het geval is) is afgenomen in plaats van gestegen ondanks de gunstige economische situatie (zie Wennekers e.a. 2019.). Naar deze bevinding is verder onderzoek nodig.

Tabel 3.1.8: Gemiddelde scores welbevinden gecorrigeerd voor compositieverschillen tussen treatmentgroepen (0-10)^{1,2}, 2017-2019

	Nulmeting	N	Meting 2	N	Meting 3	N	Totaal	N=
Eigen regie	6.00	143	6.11	77	6.14	83	6.10	303
Intensieve begeleiding	6.07	175	6.11	110	6.23	105	6.12	390
Standaardgroep	6.00	165	6.09	88	6.19	75	6.07	328
Totaal	6.02	483	6.10	275	6.19*	263	6.09*	1021

Noot¹: ***p<0,01; **p<0,05; *p<0,10, getest op basis van OLS regressie analyse (N=618); Model met treatment, cliëntvragenlijstmeting, consulentsnummer en de controle variabelen.

Noot²: *Pseudo R*²=0.075; Bijdrage verklaarde variantie door consulents=50%

Bron: Vragenlijsten nulmeting en vervolgmetingen productevaluatie Vertrouwensexperiment Oss, Tilburg University, 2020.

Lage welbevinden scores bij bijstandsgerechtigden

Tot slot valt op dat het niveau van welbevinden relatief laag is onder de deelnemers met een score van rond de 6; de gemiddelde score in Nederland is net onder 8 voor tevredenheid en geluk. Ook de verschillen tussen de deelnemers in de scores zijn relatief klein. De onderzochte groep van bijstandsgerechtigden scoort ruim twee punten lager dan gemiddeld in Nederland (zie SCP) hetgeen als een groot verschil kan worden gekenschetst. Eerder is geconstateerd dat de respons op de vragenlijsten is afgenomen na de nulmeting van 85% naar 60%. Dit heeft vermoedelijk te maken met een geringere motivatie om de vragenlijsten nog in te vullen na uitstroom uit de bijstand. Dit kan het beeld enigszins hebben gekleurd.

Invloed klantmanager op welbevinden

We hebben zoals gezegd ook een variabele in het model opgenomen die de ondersteuning door de klantmanager indiceert. Voegen we deze variabele toe aan het model dan stijgt de verklaarde variantie met 48% van 3,9% naar 7,53%. Verschillen in de ondersteuning door de klantmanager hebben dus invloed op het welbevinden van de deelnemer en verklaren voor een deel ook verschillen tussen de deelnemers in de verschillende treatments. De klantmanager doet er dus toe. De invloed van de treatments op welbevinden veranderen evenwel niet (behalve bij subjectieve gezondheid) indien we de klantmanager-variabele toevoegen aan het model. Dat betekent dat als we een treatment effect vinden dat dat een echt treatment effect is en geen effect van de persoon van de klantmanager.

Subjectieve gezondheid

Een tweede indicator die we gebruikt hebben is een indicator voor subjectieve gezondheid en mentale gezondheid. De indicator gebruikt de algemene vraag over subjectieve gezondheid (Wat vindt u van uw gezondheid?, met als antwoord categorieën: uitstekend, zeer goed, goed, matig, slecht) alsook de vijf vragen over mentale gezondheid. De vijf vragen over de mentale gezondheid betreffen de mate waarin men zich kalm en rustig voelt, men zich somber voelt en neerslachtig, zich droevig voelt, zich erg nerveus voelt en zich gelukkig voelt. De positief gestelde items zijn omgecodeerd zodat alle items in dezelfde (negatieve) richting wijzen. Voor 'de mate van' is gebruik gemaakt van een Likert schaal (altijd, vaak, soms, bijna nooit, nooit). Hiervan is een 10-puntsschaal geconstrueerd van 0-10 voor 'slechte' mentale

gezondheid die vervolgens weer is omgezet naar een positieve tien-puntsschaal van mentale gezondheid. De ontwikkeling van mentale gezondheid en fysieke gezondheid verliep in de tijd evenwel tegenovergesteld, de mentale gezondheid nam voor alle treatments toe terwijl de fysieke gezondheid of wel constant bleef voor de treatments (eigen regie, intensieve begeleiding) ofwel daalde (standaardgroep).

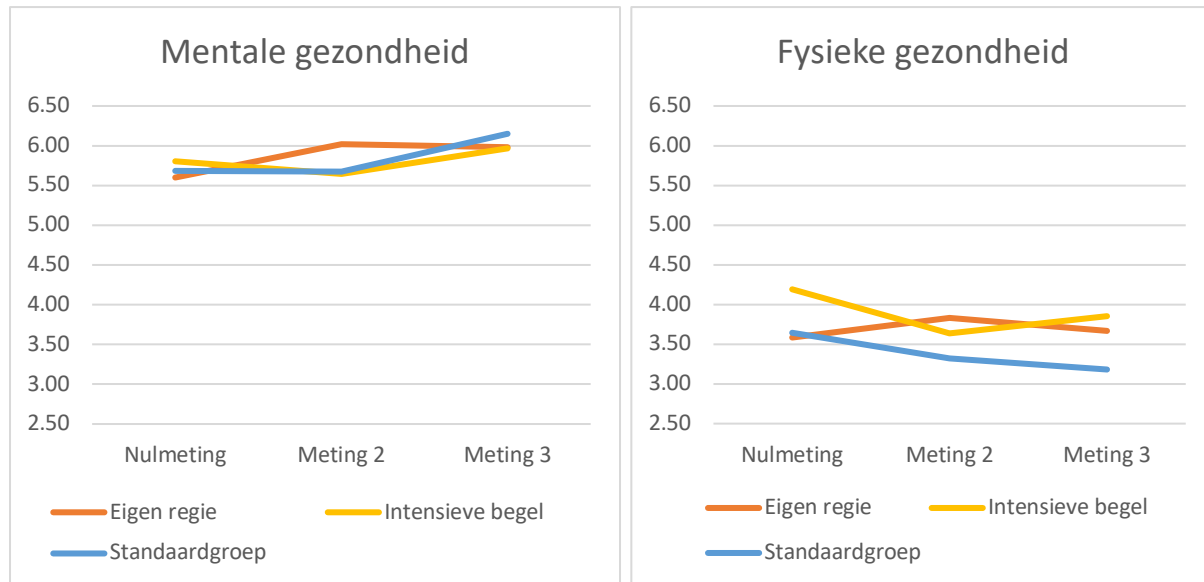


Fig. 3.1.6. Mate van fysieke en mentale gezondheid naar treatment en meting

Uit de figuur wordt duidelijk dat de fysieke en mentale gezondheidssituatie van de deelnemers te wensen overlaat. De gemiddelde score ligt bij de start rond 3.5 voor fysieke en 5.5. voor mentale gezondheid op een schaal van 0 tot 10, gemiddeld een ruime onvoldoende dus. Bij de start scoren alle groepen een ruime onvoldoende, vooral op fysieke gezondheid indien we het cijfer als een rapportcijfer voor de gezondheid opvatten. De mentale gezondheid gaat echter wel vooruit in de tijd. De stijging van de mentale gezondheid wordt significant in meting 3 maar de stijging geldt voor alle drie treatment groepen. Bij zowel fysieke als mentale gezondheid vinden we weinig tot geen treatment effecten van de beide treatments ten opzichte van de standaardgroep. De enigszins opvallende stijging van mentale gezondheid bij de standaardgroep in meting 3 is significant op 90% niveau. Dit zou erop kunnen wijzen dat de standaardgroep treatment mogelijk is veranderd. De fysieke gezondheid ontwikkelt zich slechter over de tijd gezien waarschijnlijk omdat de mensen met de relatief beste gezondheid het eerst werk vinden en dan geen vragenlijst meer invullen. De groepen gaan over de tijd bij beide uitkomstmaten iets meer uit elkaar lopen wat gezondheid betreft. De mentale gezondheid is significant toegenomen over de tijd maar wel voor alle groepen. Echter, de veronderstelling was dat de beide treatments ten opzichte van de standaardgroep een positieve bijdrage aan ervaren gezondheid zouden kunnen leveren. Deze hypothese wordt niet bevestigd maar mogelijk speelt een rol dat de standaard treatment zelf mee is veranderd. In Tabel 3.1.9 presenteren we de resultaten voor zowel mentale gezondheid als fysieke gezondheid. Uit de literatuur bleek dat intensieve begeleiding en coaching middels “Individual Placement Support” en “Employment Support”, dat zijn maatgerichte programma’s voor mensen met fysieke of mentale

gezondheidsproblemen en specifiek gericht op hun bijzondere problematiek, soelaas kan bieden om hen naar werk te begeleiden. Deze specifieke programma's zijn in dit experiment niet toegepast.

Tabel 3.1.9. Gemiddelde scores mentale en fysieke gezondheid (MGezh, FGezh) gecorrigeerd voor compositieverschillen tussen treatmentgroepen, 2017-2019

	Nulmeting		Meting 2		Meting 3		Totaal		N=
	MGezh	FGezh	MGezh	FGezh	MGezh	FGezh	MGezh	FGezh	
Eigen regie	5.60	3.58	6.02	3.83	5.98	3.67	5.81	3.67	306
Intensieve begeleiding	5.81	4.19	5.65	3.64	5.97	3.86	5.80	3.95	392
Standaardgroep	5.69	3.64	5.67	3.32	6.15*	3.18	5.79	3.45	334
Totaal	5.70	3.82	5.76	3.59	6.02**	3.60	5.80	3.70	1032

Noot: ¹⁾ Het effect is alleen significant bij het zittende bestand. ***p<0,01; **p<0,05; *p<0,10, getest op basis van OLS regressie analyse (N=); Model met treatment, cliëntvragenlijstmeting, klantmanager-nummer en de controle variabelen.

Noot²: Pseudo R²=0.057 voor MGezh en 0.162 voor FGezh ; Bijdrage verklaarde variantie door klantmanager=66% resp. 48%.

Bron: Vragenlijsten nulmeting en vervolgmetingen productevaluatie Vertrouwensexperiment Apeldoorn/Epe, Tilburg University, 2020.

Keuzevrijheid

Voor veel mensen is keuzevrijheid een belangrijk goed. De Nobelprijswinnaar Stiglitz heeft in zijn boek “De prijs van ongelijkheid” (2010) de zere vinger gelegd op het probleem van de gebrekkige ‘opportunities’ die mensen aan de onderkant van de inkomensverdeling in Amerika ervaren en hebben. Dat probleem geldt niet alleen voor de VS, ook in Nederland zijn er voor veel mensen weinig kansen en mogelijkheden om de eigen situatie te verbeteren en de dingen te doen die men voor het eigen leven belangrijk vindt. In de visie van Stiglitz is dat slecht voor het welbevinden van de mensen zelf maar ook voor de economie. Sen (1985) benadrukt dat mensen de kansen of mogelijkheden (keuzeopties) moeten hebben of krijgen om de dingen te doen in het leven die men belangrijk vindt. Het gaat zowel bij Sen als bij Stiglitz niet om gelijkheid van uitkomsten maar om gelijkheid van kansen en mogelijkheden. Om die te bieden is een ongelijke behandeling soms nodig om die gelijke startposities te kunnen realiseren terwijl in beleid veelal het gelijkheidsbeginsel leidend is. De toegang tot de arbeidsmarkt is voor mensen met beperkingen van welke aard ook beperkt en vraagt om extra ondersteuning. Het experiment is erop gericht om die kansen en mogelijkheden voor mensen aan de onderkant te bieden (maatwerk) of in ieder geval hen te leren die kansen zelf te scheppen en voor zover aanwezig te grijpen. Er is in het experiment meer vrijheid voor de klantmanager om aan de vraag van de deelnemer voor ondersteuning te voldoen of zelf met nieuwe oplossingen of instrumenten te komen niet alleen in het domein van de P-wet voorzieningen maar ook in het bredere sociale domein, zo ook in Apeldoorn/Epe. Er zijn ook minder verplichtingen voor de deelnemers van beide treatments in het experiment of beter gezegd, er wordt op een andere manier met verplichtingen omgegaan waardoor de betrokkenen meer keuzevrijheid ervaren en hebben.

In de vragenlijst is aan het aspect keuzevrijheid aandacht besteed doordat vragen zijn gesteld over de dingen die men belangrijk vindt in het leven en die men op de een of andere manier wel of niet kan doen omdat men daar de mogelijkheden wel of niet voor heeft (keuzemogelijkheden). Het zijn 7 vragen gebaseerd op Sen 's capabilities benadering (1985) die beogen inzicht te krijgen in de keuzevrijheid ofwel in de mate waarin mensen de mogelijkheden hebben om de doelen die men stelt voor het eigen leven en die men belangrijk en waardevol vindt ook daadwerkelijk in het werk of anderszins te realiseren. De lijst zou een inzicht moeten geven in de keuzemogelijkheden en beperkingen die mensen ervaren in het leven in de bijstand. In de vragenlijst zijn zeven items voorgelegd aan de deelnemers. Ten eerste wordt gevraagd of ze deze items belangrijk vinden en ten tweede of ze de mogelijkheden hebben om dit te kunnen doen in hun eigen situatie. De zeven items zijn:

- Meebeslissen over belangrijke dingen in werk en leven
- Goede contacten met anderen hebben
- Eigen doelen stellen
- Een voldoende inkomen hebben
- Iets waardevols bij te dragen aan het leven van anderen
- Om te kunnen leren en nieuwe dingen te doen (kennis en vaardigheden ontwikkelen en gebruiken)
- Om iets waardevols bij te dragen aan het leven van anderen.

Gevraagd wordt op een vijfpunts Likert schaal naar de mate waarin men dit belangrijk vindt en in de eigen situatie ook kan doen: helemaal niet, niet, soms wel, soms niet, vaak of altijd.

Capabilities indicator (keuzemogelijkheden)

In overleg met de andere onderzoekers in LOEP is een soort capabilities indicator ontwikkeld die aangeeft in welke mate mensen in staat zijn, dus keuzemogelijkheden hebben, om dingen die men belangrijk vindt te realiseren. Daarbij zijn twee varianten onderscheiden. De eerste variant berekent een score op basis van uitsluitend de items die men vaak of altijd belangrijk vindt. Items die men niet of minder belangrijk vindt worden niet meegenomen en krijgen een gewicht 0. De tweede variant weegt de score op de mate dat men deze zeven dingen kan doen met de mate waarin ze belangrijk worden geacht. Daardoor worden alle items meegenomen in de score, ook items die men minder belangrijk vindt. Deze laatste krijgen dan wel een lager gewicht. De antwoord categorieën (helemaal niet tot altijd) worden omgezet in een score van 0 tot 10. Vervolgens wordt een somscore bepaald van alle items gewogen met de mate van belangrijkheid (score 1=helemaal niet tot 5=altijd). De laatste score wordt gedeeld door 5 om gewichten te krijgen variërend van 0,2 tot 1. Deze tweede variant wordt hier gebruikt.

In Fig. 3.1.7 worden de resultaten weergegeven. In vergelijking met de standaardgroep liggen de geschatte scores voor beide treatments tot aan meting 3 iets hoger, met name bij de intensieve begeleidingsgroep maar de verschillen zijn na correctie niet significant. Bovendien zien we bij alle groepen een dalende trend hetgeen in onze visie veroorzaakt wordt doordat mensen die voltijds werk vinden en waarvan de uitkering stopt minder genegen zijn de vragenlijst in te vullen terwijl zij wellicht de meeste keuzemogelijkheden hebben. Wanneer we de uitstroom naar werk als covariaat opnemen heeft deze een significant positief effect op keuzemogelijkheden. Het kan ook zijn dat de treatments ertoe leiden dat men meer inzicht krijgt in de beperkte mogelijkheden die men heeft.

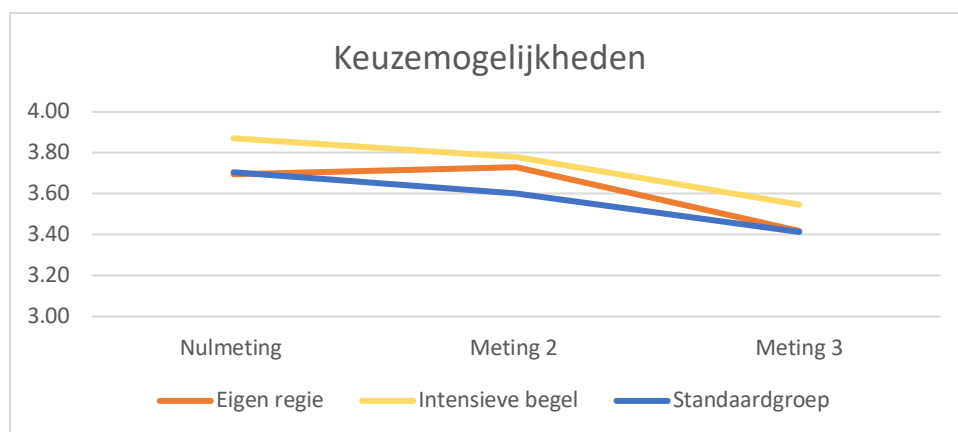


Fig. 3.1.7. Gecorrigeerde keuzemogelijkheden naar treatment en meting

Opvallend is dat de eigen regie groep eerst een lichte stijging laat zien en vervolgens een daling vanaf meting 2. De toe en afname voor de verschillende groepen is niet significant. Mogelijk wordt men pessimistischer over de eigen keuzemogelijkheden indien men ondanks eigen regie toch niet aan werk komt. In tabel 3.1.10 worden de geschatte effecten voor keuzemogelijkheden weergegeven. De verschillen tussen de treatments blijken bij toetsing niet significant te zijn. Ook over de tijd zijn de verschillen tussen de groepen niet significant.

Tabel 3.1.10. Gemiddelde scores keuzemogelijkheden (0-10) gecorrigeerd voor compositieverschillen tussen treatmentgroepen, 2017-2019

	Nulmeting		Meting 2		Meting 3		Totaal	N=
Eigen regie	3.69	144	3.73	77	3.42	85	3.63	306
Intensieve begeleiding	3.87	177	3.78	110	3.55	105	3.76	392
Standaardgroep	3.71	168	3.60	90	3.41	76	3.61	334
Totaal	3.76	489	3.71	277	3.47	266	3.67	1032

Noot¹: ***p<0,01; **p<0,05; *p<0,10, getest op basis van OLS regressie analyse (N=618); Model met treatment, cliëntvragenlijstmeting, klantmanager-nummer en de controle variabelen.

Noot²: Pseudo R²=0.091; Bijdrage verklaarde variantie door klantmanager=39%.

Bron: Vragenlijsten nulmeting en vervolgmetingen productevaluatie Vertrouwensexperiment Apeldoorn/Epe, Tilburg University, 2020.

Het niveau van keuzemogelijkheden is gemiddeld bij alle treatmentgroepen laag te noemen met een gemiddeld niveau van 3,7 op een schaal van 0 tot 10. Qua keuzemogelijkheden scoren de deelnemende bijstandsgerechtigden dus een ruime onvoldoende. Nadere analyse laat zien dat vooral op het ontbreken van mogelijkheden om de kennis en vaardigheden te gebruiken en een voldoende inkomen te hebben bij deze bijstandsgroepen laag wordt gescoord. Om te toetsen of de veronderstelling juist is dat selectieve respons de daling na meting 2 bij de beide treatments verklaart hebben we een interactie tussen uitstroom naar werk en de treatments als verklarende variabele toegevoegd. Dan blijkt dat deze een positief en significant effect heeft op de ervaren keuzemogelijkheden maar alleen bij de intensieve

begeleidingsgroep in vergelijking dus met de standaardgroep. Dat zou kunnen verklaren dat de ervaren keuzemogelijkheden afnemen ook bij intensieve begeleiding maar minder dan bij de standaardgroep het geval is. Het betekent wel dat intensieve begeleiding ertoe leidt dat men voor zichzelf meer mogelijkheden ziet dan anders het geval was geweest. In dit model met uitstroom naar werk als controle vinden we een negatief effect voor de ontwikkeling van keuzemogelijkheden in de tijd in meting 3 die dus geldt voor alle groepen.

Sociale participatie

De bevordering van sociale participatie is voor veel gemeenten een belangrijk doel van het experiment. Zelfs indien de uitstroom naar betaald werk niet haalbaar is voor de deelnemer vanwege zijn of haar problemen en beperkingen, kunnen het experiment en de treatments er wel toe bijdragen dat de sociale participatie van de betrokkenen wordt gestimuleerd. Uit de literatuur over geluk en welbevinden is bekend dat sociale participatie van belang is om verschillen in geluk tussen mensen en veranderingen in geluk van mensen te kunnen verklaren (zie Wenneker e.a., 2019; Headey et al. 2010)

In de vragenlijsten zijn daartoe vragen gesteld over de mate waarin men zich opgenomen voelt in de samenleving dus de perceptie van sociale integratie in de samenleving. Daarnaast zijn vragen gesteld over de sociale participatie, dat is het aantal uren per week dat men besteed aan vrijwilligerswerk en mantelzorg, en er zijn vragen gesteld over het sociale netwerk door te vragen naar het aantal keer per maand dat men contact heeft met buurtgenoten, kennissen en vrienden en familie. Deze vragen zijn nu gebruikt om sociale participatie indicatoren te construeren.

- *Perceptie sociale integratie:* De eerste gaat uit van de beleving van de bijstandsgerechtigde deelnemer over de mate waarin hij of zij zich welkom en opgenomen voelt in de samenleving;
- *Sociale participatie:* De tweede baseert zich op de uren besteed aan vrijwilligerswerk en mantelzorg (maatstaf voor sociale participatie) en het aantal keren dat men per maand contact heeft met anderen (buren, kennissen-vrienden, familie ofwel het sociale netwerk). De laatste noemen we het sociale netwerk. We zullen beiden afzonderlijk bekijken.

De gemeente acht de bevordering van de sociale participatie een belangrijk doel van het experiment. Voor een belangrijk deel van de cliënten is dit ook het hoogst bereikbare doel vanwege de grote afstand tot de arbeidsmarkt als gevolg van gezondheidsproblemen of een lage opleiding. Meer dan 75% heeft geen startkwalificatie en meer dan 60% heeft fysieke of mentale gezondheidsproblemen zoals uit de eerste tussenrapportage bleek (zie Muffels et al. 2018). In de procesevaluatie later in dit rapport zullen we de verandering op de participatieladder ook bekijken zoals deze door de consultants gedurende de experimenteerperiode is geregistreerd in de administratie.

Eerst wordt de beleving of perceptie van sociale integratie onderzocht, ofwel de mate waarin men zich opgenomen voelt in de samenleving en hoe deze zich ontwikkeld heeft tijdens het experiment (Fig. 3.1.8 links). Tussen meting 1 en 2 is deze licht gedaald voor alle groepen. Na meting 2 neemt de ervaren sociale integratie echter weer toe bij alle groepen. De verschillen over de tijd en tussen de afzonderlijke treatmentgroepen en de standaardgroep zijn evenwel niet significant.

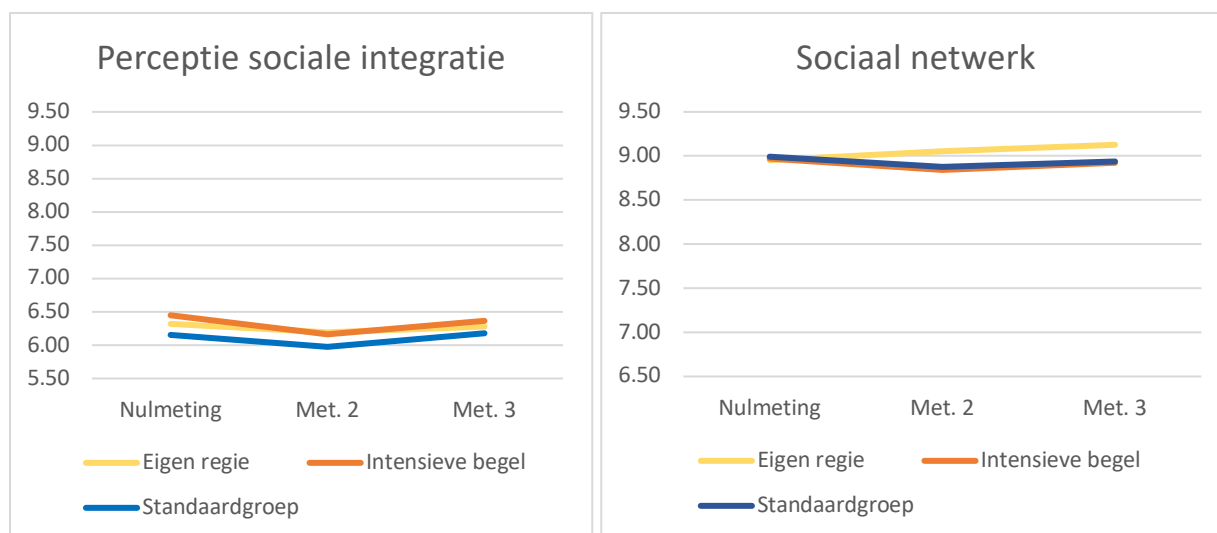


Fig. 3.1.8. Gecorrigeerde indicator perceptie sociale integratie (links) en participatie sociale netwerken (rechts) naar treatment en meting

In Fig. 3.1.8 (rechts) worden de effecten voor het sociale netwerk weergegeven. In Tabel 3.1.11 worden de resultaten voor beide uitkomstmaten weergegeven.

Tabel 3.1.11. Gemiddelden geschatte kansen perceptie sociale integratie en sociale netwerken (0-10) gecorrigeerd voor compositieverschillen tussen treatmentgroepen, 2017-2019

	Perceptie sociale integratie				Sociale netwerken				N=
	Nul-meting	Met. 2	Met. 3	Tot.	Nul-meting	Met. 2	Met. 3	Tot.	
Eigen regie	6.32	6.19	6.28	6.28	8.95	9.05	9.13	9.03	306
Intensieve begeleiding	6.45	6.16	6.36	6.35	8.97	8.84	8.92	8.92	392
Standaardgroep	6.16	5.97	6.18	6.11	8.99	8.88	8.93	8.95	334
Totaal	6.31	6.11	6.28	6.25	8.97	8.91	8.99	8.96	1032

Noot¹: ***p<0,01; **p<0,05; *p<0,10, getest op basis van OLS regressie analyse (N=1032); Model met treatment, cliëntvragenlijstmeting, klantmanagernummer en de controle variabelen.

Noot²: Pseudo R²=0.068 (links);0.048 (rechts). Bijdrage verklaarde variantie door klantmanager=40%;21%

Bron: Vragenlijsten nulmeting en vervolgmetingen productevaluatie Vertrouwensexperiment Apeldoorn/Epe, Tilburg University, 2020

We vinden ook geen significante treatment effecten. Aanvankelijk is bij intensieve begeleiding en de standaardgroep een daling waarneembaar maar bij eigen regie een stijging. Na meting 2 nemen de sociale netwerk activiteiten bij de eerste twee groepen weer toe en blijft deze bij eigen regie stijgen. Deze af of toename is niet significant en verschilt ook niet tussen groepen.

In Fig. 3.1.9 worden de resultaten voor vrijwilligerswerk en mantelzorg weergegeven.

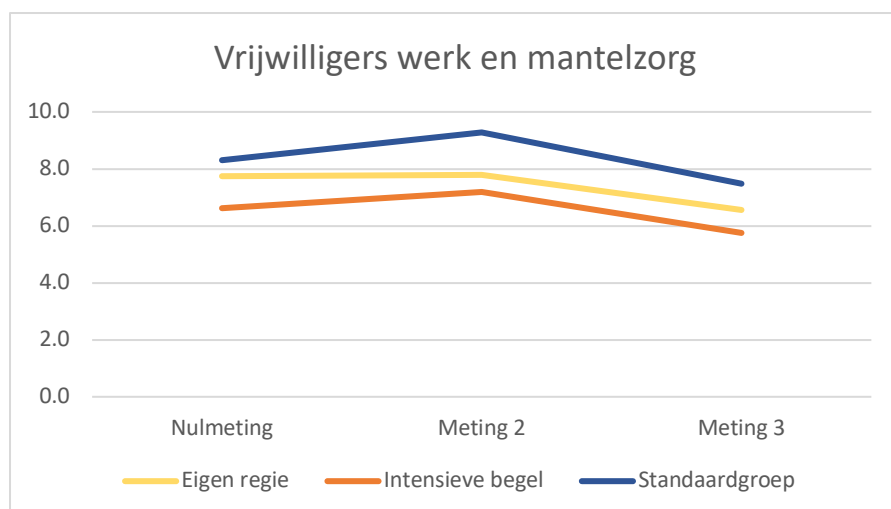


Fig. 3.1.9. Gecorrigeerde indicator vrijwilligerswerk en mantelzorg naar treatment en meting

De effecten worden weergegeven in Tabel 3.1.12. Wat dit aspect betreft van sociale participatie vinden we opnieuw eerst een toename bij intensief en de standaardgroep en vervolgens een afname bij alle groepen maar ook deze af- en toename is niet significant. We vinden twee treatment effecten; een positief effect voor het zittende bestand en een negatief effect bij intensieve begeleiding. Indien we controleren voor nieuwe instroom of zittende bestand in deze treatment dan zien we dat het effect zich voordoet bij het zittende bestand.

Tabel 3.1.12. Geschatte kansen vrijwilligerswerk en mantelzorg (0-10) gecorrigeerd voor compositieverschillen tussen treatmentgroepen, 2017-2019

	Nulmeting	Meting 2	Meting 3	Totaal	N=
Eigen regie	7.7	7.8	6.6	7.4	306
Intensieve begeleiding	6.6	7.2	5.8	6.5*	392
Standaardgroep	8.3	9.3	7.5	8.4	334
Totaal	7.5	8.0	6.5	7.4	1032

Noot¹: *** $p < 0,01$; ** $p < 0,05$; * $p < 0,10$, getest op basis van OLS regressie analyse (N=1032); Model met treatment, cliëntvragenlijstmeting, klantmanager-nummer, nieuwe instroom/zittend bestand, en de controle variabelen.

Noot²: $Pseudo R^2 = 0.048$; Bijdrage verklaarde variantie door klantmanager=31%

Bron: Vragenlijsten nulmeting en vervolgmetingen productevaluatie Vertrouwensexperiment Apeldoorn/Epe, Tilburg University, 2020

Zelfredzaamheid

De treatments zijn in belangrijke mate gestoeld op het idee dat het bevorderen van de zelfredzaamheid van de deelnemers uiteindelijk een belangrijk middel is om sociale participatie te verbeteren. De training die door Roland Blonk is aangeboden is gebaseerd op de theorie van gepland gedrag van Fishbein & Ajzen (2010). In de training van de klantmanagers is door Blonk (2019) een zoekgedrag model gebruikt welke gebaseerd is op die theorie en die als basis diende voor de zogenaamde 'skills' methodiek waarin

coaches/klantmanagers vaardigheden leren om de zelfredzaamheid van de deelnemers te bevorderen. Dat had niet alleen betrekking op de 'eigen regie' treatment maar ook op de 'intensieve begeleiding' treatment hoewel de wijze van coaching in deze laatste treatment anders, wellicht wat 'meer sturend' is als in de eerste. Maar ook in de eigen regie treatment wordt de deelnemer aangesproken op zijn of haar activiteiten en gedrag. In de vragenlijst is aan de mate van zelfredzaamheid daarom ook aandacht besteed. Er wordt onder andere gevraagd aan de deelnemer of men zichzelf zelfredzaam acht in verband met het zoeken en vinden van werk bijvoorbeeld. De lijst van capabilities eerder behandeld geeft ook inzicht in de belemmeringen en kansen om zelfredzaam te zijn, dat wil zeggen de dingen te kunnen doen die men belangrijk vindt voor het eigen leven. Wat zelfredzaamheid betreft op het gebied van werk zijn in de vragenlijst een viertal vragen gesteld daarover. Dat zijn:

- Ik kan betaald werk vinden als ik er echt moeite voor doe
- Ik denk dat ik in de toekomst betaald werk zal vinden
- Ik denk dat ik bij sollicitaties een goede indruk kan maken
- Ik denk dat ik een baan kan vinden die past bij mijn opleiding en ervaring.

De antwoordcategorieën vormen een 5-punts Likert schaal van 'geheel mee oneens' tot 'geheel mee eens'. Deze vier items zijn vervolgens omgezet in een schaal van 0 (geheel oneens) tot 10 (geheel eens) en er is een gemiddelde bepaald over de vier items. Dat noemen we de indicator zelfredzaamheid. Vervolgens is het standaard regressiemodel gedraaid om de voor verschillen in samenstelling gecorrigeerde kansen op zelfredzaamheid voor elk van de groepen te bepalen. In Figuur 3.1.10 worden de resultaten in beeld gebracht.

Treatment effecten

De zelfredzaamheid neemt geleidelijk af voor alle treatment groepen maar de afname is niet significant over de tijd gezien. De afname is evenwel het sterkst bij de standaardgroep. De afname kan verband houden met het feit dat degenen met de grootste afstand tot de arbeidsmarkt overblijven. Zij worden zich door de intensievere begeleiding meer bewust van hun eigen (minder gunstige) positie op de arbeidsmarkt. In vergelijking met de standaardgroep vinden we significante treatment verschillen voor eigen regie en intensieve begeleiding in alle drie metingen. We vinden ook een positief effect op zelfredzaamheid bij eigen regie in Epe ten opzichte van de standaardgroep in Epe (2.02; $p < 0.05$). Over alle groepen tezamen is het niveau van zelfredzaamheid onder de deelnemers in Epe echter significant lager vergeleken met Apeldoorn. Voor de beide treatment groepen ligt het niveau van zelfredzaamheid dus hoger dan voor de standaardgroep waarvoor het niveau over de gehele periode van 2 jaar daalt. Het verschil tussen beide treatment groepen in vergelijking met de standaardgroep is significant zo blijkt uit de regressieanalyse (zie Tabel 3.1.13).

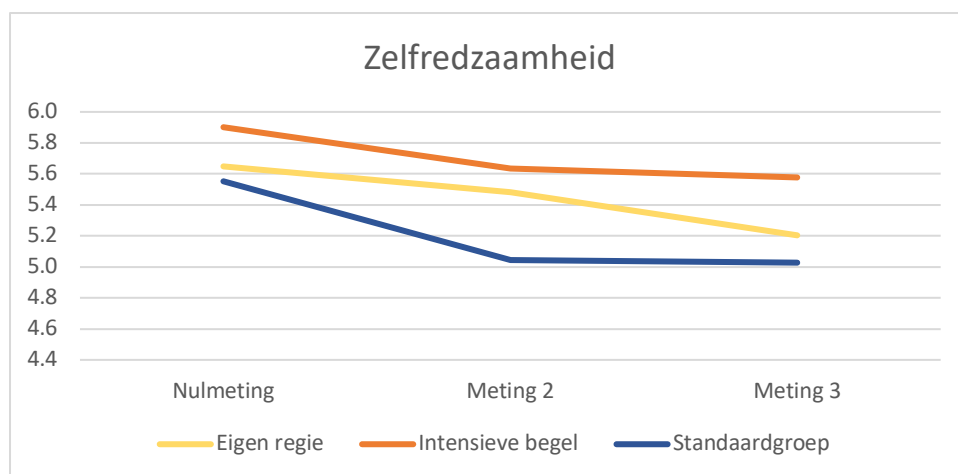


Fig. 3.1.10. Gecorrigeerde indicator zelfredzaamheid bij zoeken naar en vinden van werk naar treatment en meting

Zoals gezegd kan selectie in meting 3 de daling over de tijd hebben beïnvloed aangezien op dat moment de groep met de kleinste kansen op werk is overgebleven en de mensen die zijn uitgestroomd de vragenlijst niet hebben ingevuld. Voegen we uitstroom naar werk toe aan de regressie dan is het effect positief en significant op 95% niveau. Causaal gezien leidt zelfredzaamheid tot uitstroom en niet andersom hetgeen betekent dat de meest zelfredzame deelnemers het eerst uitstromen. In interactie met de treatments is het effect uitsluitend significant bij de eigen regie groep waarin zelfredzaamheid het uitgangspunt van de treatment is. De meest zelfredzame mensen stromen dus het eerst uit maar in vergelijking met de standaardgroep vooral bij de eigen regiegroep. We vinden voorts een negatief hoofdeffect van zittend bestand (-0.38; $p < 0.05$). Tabel 3.1.13 geeft het gemiddelde van de geschatte kansen per treatment en meting en de significantieniveaus.

Tabel 3.1.13 Gemiddelde scores zelfredzaamheid (0-10) gecorrigeerd voor compositieverschillen tussen treatmentgroepen, 2017-2019

	Nulmeting	Meting 2	Meting 3	Totaal	N=
Eigen regie	5.6*	5.5*	5.2*	5.5***	306
Intensieve begeleiding	5.9	5.6	5.6	5.7*	392
Standaardgroep	5.6	5.0	5.0	5.3	334
Totaal	5.7	5.4	5.3	5.5	1032

Noot¹: *** $p < 0,01$; ** $p < 0,05$; * $p < 0,10$, getest op basis van OLS regressie analyse $N=618$); Model met treatment, cliëntvragenlijstmeting, klantmanagernummer en de controle variabelen.

Noot²: $Pseudo R^2=0.222$; Bijdrage verklaarde variantie door klantmanager=35%

Bron: Vragenlijsten nulmeting en vervolgmetingen productevaluatie Vertrouwensexperiment Apeldoorn/Epe, Tilburg University, 2020.

Vertrouwen

Het experiment is vertrouwensexperiment genoemd omdat het vertrekt vanuit het idee dat het opbouwen van een vertrouwensrelatie met de deelnemer uiteindelijk effectiever kan zijn om het gewenste doel te bereiken, namelijk bevordering van sociale participatie, dan de standaardaanpak met een strenge toepassing van de verplichtingen en bij niet nakoming sanctionering van de uitkering. De toetsing van rechtmatigheid blijft intact maar de toetsing aan de andere (informatie, sollicitatie, tegenprestatie) verplichtingen wordt op een meer relaxte manier vormgegeven in nauwe afstemming en overleg met de klant. Het is zeker geen 'loslaten' maar leren de eigen verantwoordelijkheid te nemen en zelfredzaam te zijn met het aanpakken van de problemen die men heeft en ervaart. Uit de experimentele economische literatuur weten we bijvoorbeeld dat het geven van vertrouwen leidt tot wederkerigheid, tot extra inspanning en verbeterde motivatie en productiviteit in reactie op het gestelde vertrouwen (Frey & Jegen 2001; Bohnet et al., 2001). Uit de psychologische motivatietheorie weten we dat extrinsieke motivatie ten koste kan gaan van de intrinsieke motivatie en dat sturing op intrinsieke motivatie beter kan werken. Vandaar ook de inbouw van de training van de klantmanagers om hen vaardigheden te leren die hen helpen mensen te ondersteunen in zelfredzaamheid. Apeldoorn/Epe is als niet-AMvB gemeente niet gemachtigd ruimere vrijlatingsbepalingen voor extra verdiensten dan in de wet staat toe te staan. Dat is voorbehouden aan de artikel 83 gemeenten. Daardoor is er geen treatment die het aspect belonen in plaats van straffen kan toetsen op zijn werkbaarheid. Er vindt wel beloning plaats maar dan in de vorm van intensieve begeleiding en maatwerk, een luisterend oor en aandacht. Doordat de caseload van de klantmanager verlaagd is van 125 naar 75 tot 90, is er meer tijd voor de begeleiding van de klant. Er is tijd om een 'relatie' en of persoonlijke band met de deelnemer op te bouwen omdat er meer tijd voor contact is en meer coaching op basis van vertrouwen. Het wantrouwen onder de deelnemers ten opzichte van de gemeente is groot zo hebben alle klantmanagers ervaren. Om daar verandering in te brengen kost tijd (zie verslag Focusgroep gesprekken hierna). Vertrouwen komt te voet en gaat te paard. Dus het vertrouwen is ook weer snel weg indien het vertrouwen wordt geschaad; dat geldt overigens voor beide partijen.

Over dit aspect van vertrouwen hebben we vragen gesteld in de vragenlijst. Ten eerste betreft dat een vraag over vertrouwen in de instituties. De vraag luidde in welke mate men vertrouwen had in:

- De landelijke overheid (regering, parlement)
- De gemeente
- De afdeling werk en inkomen van de gemeente
- De eigen klantmanager/werkcoach.

De antwoord categorieën vormde een vier-puntsschaal van 'helemaal geen vertrouwen', 'niet zo veel vertrouwen' 'een beetje vertrouwen' tot 'veel vertrouwen'. Deze is omgezet naar een 0-10 schaal. Daarnaast is gevraagd naar het vertrouwen dat men in andere mensen heeft. Dat wordt in de literatuur aangeduid als sociaal vertrouwen. Het wordt gezien als een vorm van sociaal kapitaal, mensen met veel vertrouwen hebben veelal een rijk sociaal netwerk dat hen ook helpt om vooruit te komen in het leven. Sociaal vertrouwen of vertrouwen in andere mensen is gemeten met een vraag die afkomstig is uit het "European Social Survey" of men vindt dat de meeste mensen te vertrouwen zijn of dat men niet voorzichtig genoeg kan zijn op een schaal van 0 (geen vertrouwen) tot 10 (volledig vertrouwen).

In Fig. 3.1.11 wordt de ontwikkeling in vertrouwen volgens de twee maatstaven per treatment groep weergegeven. Het vertrouwen in de instituties en het vertrouwen in anderen is op een laag niveau maar wel stabiel. Beiden dalen licht voor twee van de drie treatments na meting 1. Echter, we vinden geen verschillen tussen de groepen en ook de afname in de tijd is niet betekenisvol. Na het eerste jaar neemt het sociaal vertrouwen af. Het niveau van sociaal vertrouwen is hoger dan het niveau van institutioneel vertrouwen. Het relatief lage niveau van institutioneel vertrouwen weerspiegelt het wantrouwen van de deelnemers ten aanzien van de overheid in het algemeen. Het is onduidelijk waarom bij de start het niveau van sociale institutioneel vertrouwen bij de standaardgroep iets lager ligt. Het verschil is echter niet betekenisvol.

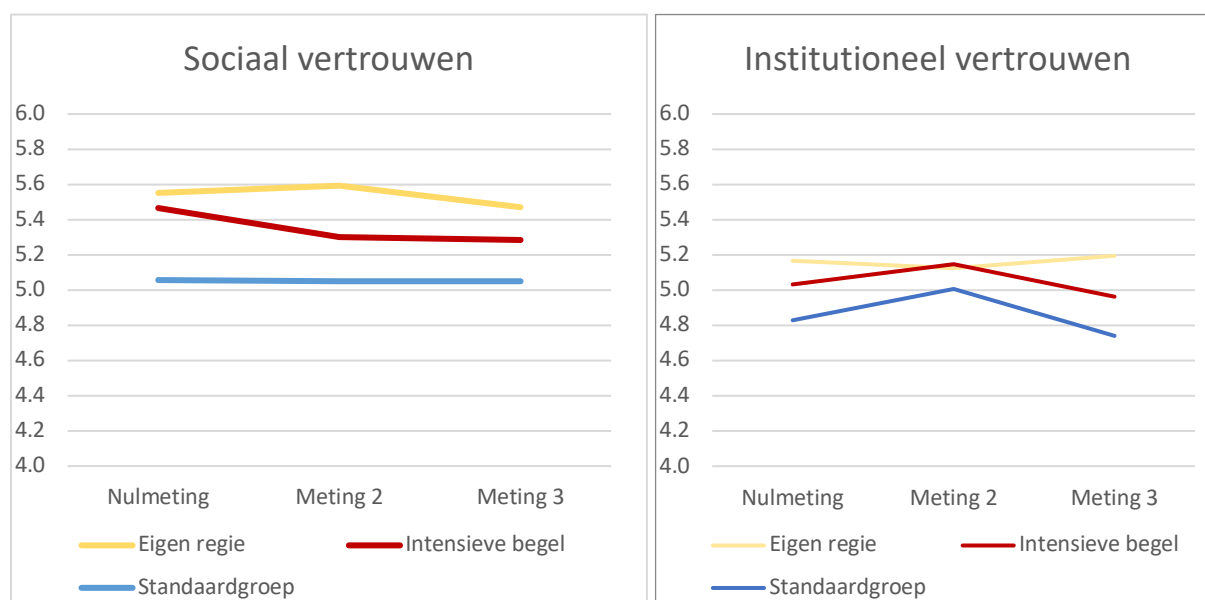


Fig. 3.1.11 Gecorrigeerde indicatoren vertrouwen in instituties en sociaal vertrouwen naar treatment en meting

In Tabel 3.1.14 worden de cijfers voor zowel institutioneel en sociaal vertrouwen in tabelvorm weergegeven. We vinden voor zowel institutioneel als sociaal vertrouwen geen causale treatment effecten in de vergelijking met de standaardgroep. Ook zijn er geen betekenisvolle verschillen tussen beide treatments. De veranderingen in institutioneel vertrouwen en sociaal vertrouwen in de tijd wijzen eveneens niet op een treatment effect.

Tabel 3.1.14 Gemiddelde scores institutioneel en sociaal vertrouwen (0-10) gecorrigeerd voor compositieverschillen tussen treatmentgroepen, 2017-2019

	Sociaal vertrouwen				N=	Institutioneel vertrouwen				N=
	Nulmt	Mt2	Mt 3	Tot.		Nulmt	Mt 2	Mt3	Tot.	
Eigen regie	5.6	5.6	5.5	5.5	288	5.2	5.1	5.2	5.2	295
Intensieve begeleiding	5.5	5.3	5.3	5.4	382	5.0	5.1	5.0	5.0	387
Standaardgroep	5.1	5.1	5.1	5.1	315	4.8	5.0	4.7	4.9	319
Totaal	5.4	5.3	5.3	5.3	985	5.0	5.1	5.0	5.0	1001

Noot¹: ***p<0,01; **p<0,05; *p<0,10, getest op basis van OLS regressie analyse N=985;1001); Model met treatment, cliëntvragenlijstmeting, klantmanagernummer en de controle variabelen.

Noot²: Pseudo R²: Intrust=0.064; Strust=0.086. Bijdrage verklaarde variantie door klantmanager ltrust=25%; Strust=26%

Bron: Vragenlijsten nulmeting en vervolgmetingen productevaluatie Vertrouwensexperiment Apeldoorn/Epe, Tilburg University, 2020

Financiële situatie en armoede

In de niet-AMvB gemeenten is het niet wettelijk mogelijk om een vrijlatingsgroep in te richten waarbij mensen een extra deel van hun verdiensten mogen behouden. Dat is voorbehouden aan de gemeenten die zijn erkend als experimenteer gemeente. Verwacht mag worden dat voor deze groepen in deze gemeenten de financiële situatie wordt verbeterd en dat dit effecten heeft op de ervaren stress die een slechte financiële situatie en armoede met zich meebrengt. Nochtans kunnen er ook in Apeldoorn/Epe als niet-AMvB gemeente effecten zijn van een andere bejegening en aanpak bijvoorbeeld doordat meer mensen voltijds werk vinden en volledig onafhankelijk worden van een uitkering of dat schuldproblemen worden aangepakt of dat mensen gestimuleerd worden om hun huishoudboekje beter bij te houden.

Een van de veronderstellingen die we op grond van de psychologische literatuur naar voren hebben gebracht is dat door ervaren financiële stress mensen cognitief en motivationeel daar zo mee bezig zijn dat er mentaal geen ruimte is voor activiteiten die de eigen situatie op iets langere termijn zouden kunnen verbeteren zoals zoeken naar werk, het volgen van een opleiding, het inschakelen van het sociale netwerk en dergelijke.

Inkomensarmoede

Er zijn drie maatstaven gebruikt om de financiële situatie en de armoedesituatie in beeld te brengen. De eerste is een maatstaf welke veel is gebruikt in onderzoek naar de inkomenssituatie van huishoudens in Nederland. Daarbij wordt direct gevraagd naar de financiële situatie van het huishouden waarin men leeft. De vraag luidt als volgt:

Wat is op dit moment de financiële situatie van uw huishouden? De antwoordcategorieën zijn:

- Ik maak schulden
- Ik gebruik mijn spaarmiddelen om van te leven
- Ik kan precies rondkomen
- Ik houd een beetje geld over

- Ik houd genoeg over om te sparen
- Ik weet het niet
- Ik wil dit niet zeggen

In de vraag wordt tevens vermeld wat het begrip rondkomen inhoudt. Rondkomen is net genoeg geld hebben om de absoluut minimaal noodzakelijke kosten van levensonderhoud te kunnen betalen.

De eerste twee items zijn nu gebruikt om een persoon inkomensarm te noemen. Er zijn in de literatuur veel armoedemaatstaven ontwikkeld en ook verschillende inkomensarmoede maatstaven. De bijstandsgrens is zo'n grens die het SCP bijvoorbeeld hanteert naast een aantal andere. Een daarvan is het "niet-veel-maar-toereikend budget" dat ongeveer op het niveau van de bijstand van een alleenstaande ligt. We verwijzen naar de armoederapporten van het SCP die regelmatig verschijnen (Hoff e.a., 2019). Ook het CBS publiceert cijfers over armoede. Het SCP heeft berekend dat in 2014 ongeveer 45% van de mensen in de bijstand onder de "niet-veel-maar-toereikend" budgetgrens leven. Recente cijfers laten zien dat dit gedaald is naar 25-35% in 2019. De maatstaf voor armoede op basis van de vraag naar de financiële situatie noemen we inkomensarmoede en geeft het percentage deelnemers dat niet kan rondkomen met het huishoudinkomen en ofwel de spaarmiddelen aanspreekt om rond te kunnen komen ofwel schulden moet maken.

Deprivatiearmoede

De tweede maatstaf die we hebben gebruikt noemen we deprivatiearmoede en is gebaseerd op twee vragen waarin gevraagd wordt bij elk van de 14 items of men deze absoluut noodzakelijk vindt om te hebben of te kunnen doen en of men voldoende geld heeft om het te kunnen kopen of te kunnen doen. De lijst van 14 items is een selectie uit een lijst van 21 items die ook is bevraagd in onderzoek van de Europese Commissie, EU-SILC genaamd. We hebben er de 14 voor de Nederlandse samenleving relevante items uit geselecteerd. Dat zijn de volgende.

- 1 keer per dag groenten en fruit
- 1 maaltijd per dag met vlees kip of vis
- Twee paar schoenen
- Laptop of desktop computer
- Vrijtijdsactiviteiten buiten het gezin
- Vieren van bijzondere gebeurtenissen
- Jaarlijks 1 week op vakantie gaan in het buitenland
- Vervangen van beschadigde meubels
- Vervangen van huishoudelijke apparaten
- Kleding vervangen
- 1 keer per maand uit eten gaan
- Regelmatig deelnemen aan vrijetijdsactiviteiten
- Toegang tot internet
- Gebruik kunnen maken van openbaar vervoer

Vervolgens is het aantal items berekend dat men noodzakelijk vindt maar niet heeft. Voor de maatstaf van deprivatiearmoede zijn er vervolgens vijf essentiële items geselecteerd: 1x per dag groenten en fruit; 1 maaltijd per dag met vlees, kip of vis; vervangen beschadigde meubels; vervangen van huishoudelijke

apparaten en versleten kleding vervangen. Indien iemand elk van deze items noodzakelijk vindt maar er minstens drie van deze vijf items mist beschouwen we deze persoon als arm. We noemen iemand dan deprivatiearm.

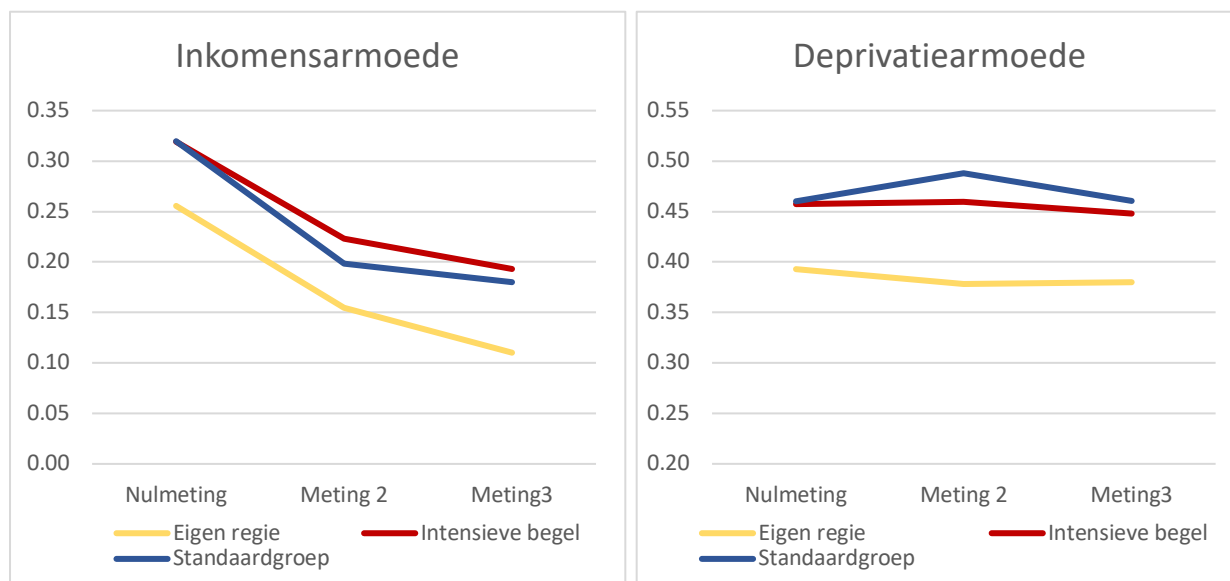


Fig. 3.1.12 Gecorrigeerde indicatoren voor financiële stress en deprivatiearmoede

Naast 20 tot 25% die in inkomensarmoede leeft, leeft zo'n 40 tot 50% van de deelnemers in deprivatiearmoede. Zeker ook dit laatste cijfer is verontrustend. Niettemin laat Fig. 3.1.12 een daling zien van inkomensarmoede in de tijd die zowel voor meting 2 als 3 significant is. Deprivatiearmoede kent evenwel een stabielere ontwikkeling. Voorts vinden we een positief en significant effect voor het zittende bestand in alle groepen hetgeen betekent dat de kans op deprivatiearmoede in het zittende bestand het grootst is. De daling in inkomensarmoede kan het gevolg zijn van de gunstige economische situatie maar het kan ook zijn dat de uitstroom naar werk (20%) in Apeldoorn/Epe hier debet aan is geweest. Echter, zoals we eerder constateerden is deze uitstroomwinst niet beperkt tot de beide experimentgroepen maar realiseerde ook de standaardgroep gunstige uitstroomcijfers die enkele procentpunten hoger waren dan de uitstroomcijfers van de referentiegroep van niet-deelnemers. Controleren we bij deze treatment voor uitstroom naar werk dan heeft dit geen effect op de inkomens -of deprivatiearmoede bij de groepen. Kennelijk zijn er andere factoren die het niveau van inkomensarmoede en deprivatiearmoede verklaren (zoals het inkomensbeleid).

Treatment effecten

De grafiek laat ook zien dat de inkomensarmoede over de tijd sterk is gedaald vooral bij de eigen regiegroep. Deze daling is significant ten opzichte van de standaardgroep. We vinden tevens significante treatment effecten voor zowel eigen regie als intensieve begeleiding. Opvallend is dat eigen regie in Epe een sterk negatief effect laat zien, hetgeen betekent dat de deprivatiearmoede in deze groep significant is verminderd. Opvallend is nog wel dat de beide maatstaven iets anders lijken te meten. De standaardgroep heeft de hoogste deprivatiearmoede maar niet de hoogste inkomensarmoede. De eigen

regiegroep heeft wel de laagste deprivatiearmoede en inkomensarmoede. De correlatie tussen beide is laag (0,10) hetgeen bevestigt dat ze iets anders meten. De inkomensarmoede meet vooral de perceptie van de eigen situatie en uit de armoedeliteratuur weten we dat mensen zich aanpassen en leren leven met het inkomen of de dingen die men heeft.

Tabel 3.1.14 Gemiddelde scores inkomensarmoede en deprivatiearmoede (0-10) gecorrigeerd voor compositieverschillen tussen treatmentgroepen, 2017-2019

	Deprivatiearmoede				Inkomensarmoede				N=
	Nulmt	Mt2	Mt 3	Tot.	Nulmt	Mt 2	Mt3	Totaal	
Eigen regie	0.39	0.38	0.38	0.38***	0.26	0.15	0.11	0.19*	0.39
Intensieve begeleiding	0.46	0.46	0.45	0.46**	0.32	0.22	0.19	0.26	0.46
Standaard	0.46	0.49	0.46	0.47	0.32	0.20	0.18	0.26	0.46
Totaal	0.44	0.45	0.43	0.44	0.30	0.20***	0.16***	0.24	0.44

Noot¹: ***p<0,01; **p<0,05; *p<0,10, getest op basis van OLS regressie analyse N=618); Model met treatment, cliëntvragenlijstmeting, klantmanagernummer en de controle variabelen.

Noot²: *Pseudo R²: Ink. armoede. armoede 0,098; Depriv.=0.10.* Bijdrage verklaarde variantie door klantmanager 13%;35%

Bron: Vragenlijsten nulmeting en vervolgmetingen productevaluatie Vertrouwensexperiment Apeldoorn/Epe, Tilburg University, 2020

Samenvatting en conclusies: treatment effecten

In Tabel A1.1 in bijlage A presenteren we een overzicht van de gevonden treatment effecten ten opzichte van de standaardgroep en de effecten van de beide metingen voor alle drie groepen ten opzichte van de nulmeting op de niet-werk gerelateerde uitkomstmaten (ontwikkeling in de tijd). De meeste treatment effecten zijn niet significant. De meeste hypothesen worden daarom niet bevestigd met uitzondering van de verwachtingen ten aanzien van zelfredzaamheid en armoede. Daar vinden we positieve treatment effecten voor eigen regie en intensieve begeleiding vergeleken met de standaardgroep. De effecten zijn betrekkelijk groot zo blijkt uit tabel A1.1. Dit zijn relevante bevindingen in het licht van de kwetsbaarheid van de doelgroep in Apeldoorn/Epe waar drie van de vier deelnemers een serieus gezondheidsprobleem ervaart dan wel zodanige beperkingen heeft dat werk voor hen betaald werk een verre einder is zeker indien eveneens drie van de vier niet de minimum kwalificaties heeft om werk te vinden in de laagstbetaalde banen.

Bij welbevinden vinden we geen treatment effect ten opzichte van de standaardgroep maar stijgt het welbevinden nochtans bij alle groepen betekenisvol. Datzelfde geldt ook voor mentale gezondheid. We vinden een betekenisvolle stijging in de tijd bij alle groepen. Dat er weinig treatmenteffecten gevonden worden bij deze uitkomstmaten heeft volgens ons te maken met twee aspecten. Ten eerste denken we dat we geen goede vergelijkingsgroep hebben omdat de standaard treatment zelf is veranderd. Ten tweede denken we dat veel van de gekozen uitkomstmaten stabiele kenmerken zijn van mensen die niet zo snel veranderen, of het nu gaat om beleving en motivatie zoals bij welbevinden, subjectieve gezondheid, de perceptie van sociale integratie, vertrouwen of om gedrag zoals bij zoekgedrag,

vrijwilligerswerk en sociale netwerken. Om deze kenmerken te veranderen zijn vaak ingrijpende gebeurtenissen nodig die het leven drastisch veranderen zoals een scheiding, langdurige werkloosheid, en verlies van kind partner of vriend en vaak duurt het een lange tijd voor een effect zich openbaart door een andere beleving, motivatie of gedrag. Wat het eerste aspect betreft laten de grafieken ook zien dat de uitkomsten voor de standaardgroep niet stabiel zijn en in dezelfde richting mee veranderen met de andere groepen. Bij de uitkomstmaat van mentale gezondheid vinden we bijvoorbeeld een positief significant treatment effect over de tijd bij de standaardgroep in meting 3 (zie Tabel A1.2 in Bijlage A). Indien de standaard treatment niet was veranderd zou het vinden van een dergelijk effect onwaarschijnlijk zijn geweest. Daarnaast zagen we al op grond van de analyses van de uitstroom naar werk dat de standaard treatment naar alle waarschijnlijkheid is veranderd omdat er anders niet zo'n groot verschil zou zijn met de referentiegroep die de standaard treatment heeft gekregen. Met andere woorden, indien de standaard treatment inderdaad is veranderd, is geen zuivere vergelijking met de treatment groepen meer mogelijk. Dan zullen we een vergelijking moeten maken met een vergelijkingsgroep die niet beïnvloed is door het experiment. Op dit moment is zo'n alternatieve vergelijkingsgroep niet beschikbaar. Daarvoor zullen we te rade moeten gaan bij andere databronnen.

Effecten over tijd

We vinden dus wel positieve en significante effecten over tijd in verband met welbevinden en mentale gezondheid maar dan voor alle groepen. Of dit het gevolg is van het experiment, dus een treatment effect, is vooralsnog causaal niet vast te stellen zonder zo'n vergelijkingsgroep. Het is wel opvallend omdat er geen evidentie is dat welbevinden bij bijstandsgerechtigden zou zijn toegenomen in de afgelopen twee jaar, integendeel, cijfers van het SCP suggereren (zie Van Echteld, 2019) dat voor mensen met een levenssatisfactiescore van 6 of lager (waartoe ook de bijstandsgerechtigden behoren) de score de afgelopen jaren is verslechterd. Het is dus aannemelijk om te veronderstellen dat we hier te maken hebben met een 'Hawthorne' effect bij de standaardgroep maar zeker weten doen we het niet en vraagt nader onderzoek.

Negatieve treatment effecten

Een uitkomst die benadrukt moet worden is dat we nauwelijks negatieve effecten vinden. Dat betekent dat de andere werkwijze in ieder geval niet slechter werkt op deze uitkomstmaten dan de klassieke methode. We vinden slechts één negatief effect. Dat is het geval bij vrijwilligerswerk en mantelzorg. Het betreft in dit geval de intensieve begeleidingsgroep in vergelijking met de standaardgroep. Gezien het relatief grote aantal afmeldingen in deze groep komend uit het zittend bestand heeft er enige zelfselectie plaatsgevonden. In hoeverre ook selectieve respons in de vragenlijsten bestaat vraagt om nader onderzoek.

Kortom...

We vinden enkele interessante treatment effecten zoals bij zelfredzaamheid en deprivatiearmoede maar anderen worden ogenschijnlijk weerlegd. Toch ontwikkelen een aantal belangrijke uitkomstmaten zich in de tijd gunstig maar niet beter dan in de standaardgroep. We vinden geen causaal verband met de treatments met uitzondering van de genoemde. Nader onderzoek moet uitwijzen of onderliggend bijvoorbeeld de stijging in welbevinden en mentale gezondheid niet toch een treatment effect is dat niet

zichtbaar wordt omdat de standaard treatment gedurende het experiment is veranderd en is gaan lijken op de andere treatments. We vinden ook verschillen tussen eigen regie en intensieve begeleiding. Veelal zijn de scores van eigen regie beter maar hier speelt mogelijk het selectie effect bij intensieve begeleiding bij de aanmeldingen een rol.

Het gevonden positieve treatment effect op zelfredzaamheid en armoede is zeker niet onbelangrijk aangezien zelfredzaamheid een van de middelen was die de gemeente zich bij aanvang stelde. Eerder constateerden we al dat er positieve effecten zijn op uitstroom naar werk en bij deze uitkomstmaat konden we controleren voor het 'Hawthorne' effect. Daar vonden we sterke aanwijzingen voor zo'n effect, maar wel een die het onderzoek heeft bemoeilijkt bij de niet-werk gerelateerde uitkomstmaten.

Hierna bespreken we de uitkomsten van de procesevaluatie en analyseren we het uitvoeringsproces zelf aan de hand van de vragenlijsten gesteld aan de klantmanager. Wellicht krijgen we hiermee meer inzicht in de wijze waarop de standaard treatment zich heeft ontwikkeld.

3.2 Procesevaluatie: vragenlijsten voor klantmanagers

De procesevaluatievragenlijsten, zijn de vragenlijsten die door de betreffende klantmanager zijn ingevuld voor de deelnemers die deze klantmanager in begeleiding heeft. Een klantmanager heeft dus voor al zijn of haar cliënten een of meerdere keren de vragenlijst ingevuld. Deze vragenlijst is zoals eerder aangegeven in totaal drie keer ingevuld: de nulmeting, de eerste vervolgmeting en de tweede vervolgmeting. De resultaten van deze vragenlijsten geven inzicht in het proces van begeleiding vanuit het perspectief van de klantmanager. Er wordt een beeld geschetst van de ontwikkeling die de deelnemer heeft doorgemaakt volgens de klantmanager en in hoeverre er sprake was van een effectieve begeleiding en of er zich leereffecten hebben voorgedaan voor zowel deelnemer als klantmanager.

3.2.1 Gehanteerde uitkomstmaten

Om een inzicht te krijgen in het proces van begeleiding en de ontwikkeling van de deelnemer door de ogen van de klantmanager, wordt er naar de volgende aspecten gekeken:

- Hoe beoordeelden de klantmanagers de **motivatie** van de deelnemers voor de treatment en de te behalen doelen en hoe veranderde dit over de tijd?
- Hoe beoordeelden de klantmanagers de deelnemer ten aanzien van een viertal aspecten van de begeleiding: of de deelnemer **realistische eisen** stelt ten aanzien van de (hoogst bereikbare) **doelen**; het **zelf in actie** komen, het **zelf doelen stellen** en het verwachte **effect van de treatment** gedurende het experiment.
- Hoe beoordeelden de klantmanagers de mate van **sociale participatie** en de **afstand tot de arbeidsmarkt** voor de deelnemer (de plaats op de participatieladder) en hoe veranderde dit over de tijd?
- Hoe beoordeelden de klantmanagers de **leereffecten** bij de deelnemer (mate van **zelfredzaamheid**) en bij zichzelf en hoe veranderde dit over de tijd?
- In hoeverre beoordeelden de klantmanagers de geboden treatment als geschikt voor de deelnemer (alleen in de derde vragenlijst bevraagd)?

Uiteraard is een belangrijk doel van het onderzoek om te bekijken in hoeverre er verschillen zijn tussen de deelnemers in de drie treatments op de bovengenoemde aspecten. Daarbij gaat het om verschillen er op een bepaald moment maar ook in hoeverre deze verschillen veranderen over de tijd en in hoeverre deze veranderingen dan zijn toe te schrijven aan het effect van de treatment.

3.2.2 Analysemethode vragenlijst

Op basis van de verzamelde data van de klantmanager vragenlijsten is er een gewone regressieanalyse (OLS) uitgevoerd, waarbij het verband tussen de treatments en de gehanteerde uitkomstmaten wordt geschat. In deze analyse zijn een aantal onafhankelijke variabelen toegevoegd om de schatting van de uitkomstmaat (afhankelijke variabele) zo zuiver mogelijk te maken, namelijk: de meting waarin de score is behaald op de uitkomstmaat, het nummer van de klantmanager/consulent die de vragenlijst heeft ingevuld en een dummy variabele die aangeeft of de deelnemer uit het zittende bestand komt, of uit de nieuwe instroom. Daarnaast is er ook gecontroleerd voor de gemeente waaruit de deelnemer afkomstig is: Apeldoorn en Epe. Later in de rapportage wordt ingegaan op de gevonden verschillen tussen deze gemeenten op de uitkomstmaten. In de rapportage van de analyse, wordt er gebruik gemaakt van een beschrijvende tabel waarin de gemiddelde scores voor de treatment groepen op een uitkomstvariabele worden getoond per meting en in totaal. Deze gemiddelden zijn geschat op basis van de eerdergenoemde regressieanalyse. Een nadere uitleg over hoe de tabellen gelezen moeten worden, volgt in de volgende paragraaf.

3.2.3 Resultaten

In deze paragraaf zullen de resultaten op de vragen uit paragraaf 3.2.1 worden beschreven. De tabellen die worden getoond zijn als volgt te lezen. In de tabellen zijn de gemiddelde scores op de uitkomstvariabele (bijvoorbeeld motivatie, plaats op participatieladder, zelfredzaamheid, etc.) per treatmentgroep en per meting te zien. Deze gemiddelden zijn *geschat* op basis van een regressieanalyse waarbij de treatment groepen en de metingen de onafhankelijke variabelen waren en de uitkomstvariabelen telkens de afhankelijke variabele. Daarnaast zijn er nog drie onafhankelijke variabelen gebruikt, namelijk: klantmanagernummer, uit welke gemeente de deelnemer komt (Apeldoorn of Epe) en een variabele die aangeeft of een deelnemer uit het zittend bestand of nieuwe instroom komt¹⁶. Er kan dus gezegd worden dat de gemiddelden die hier zijn weergegeven, de gemiddelden zijn onafhankelijk van (of gecorrigeerd voor) het effect van de klantmanager, uit welke gemeente iemand komt en of iemand afkomstig is uit het zittend bestand of nieuwe instroom.¹⁷ Hoe moeten de tabellen gelezen worden? De

¹⁶ Er zijn ook regressiemodellen getest waarin een meer uitgebreide verzameling aan onafhankelijke variabelen is opgenomen om de robuustheid van de resultaten te testen. In deze modellen werd er constant gehouden voor de volgende kenmerken van de klant: leeftijd, geslacht, opleidingsniveau, soort woning, leefvorm, duur van uitkering en nationaliteit. Interessante bevindingen uit deze analyses worden steeds bij de resultaten besproken.

¹⁷ In totaal zijn er voor elke uitkomstvariabele twee regressiemodellen geschat: een zonder interactie tussen de treatments en meting en een met de interacties. De regressie coëfficiënten (effectmetingen) voor de effecten van de treatments en de metingen van het eerste model zijn opgenomen in Bijlage B, tabel 1. De regressie coëfficiënten

kolom 'gemiddelde totaal' geeft de gemiddelde score op de 3 metingen voor elke treatmentgroep. De sterretjes geven significantieniveaus weer (zie ook Bijlage B voor bijbehorende geschatte effecten). Bij drie sterretjes weten we met 99% betrouwbaarheid dat het verschil tussen de treatments (deels)¹⁸ wordt veroorzaakt door de treatment zelf. Bij twee sterretjes weten we dit met 95% betrouwbaarheid en bij een ster met 90% betrouwbaarheid. Ontbreken de sterretjes, dan wordt het verschil waarschijnlijk door iets anders veroorzaakt dan door het effect van de treatment. De onderste rij met 'Totaal', geeft de gemiddelden over de groepen per meting, dus over de tijd. Sterretjes in deze rij laten zien dat er een significante of betekenisvolle toename of afname plaatsvindt over de tijd ten opzichte van de *nulmeting*. Kijken we naar de resultaten in de nulmeting dan kunnen we sterretjes interpreteren als betekenisvolle effecten van de treatments eigen regie en intensieve begeleiding vergeleken met de *standaard treatment* in die meting. Bij de resultaten voor vervolgmetingen 1 en 2 geven de sterretjes de effecten (een betekenisvolle toe-of afname) weer van een treatment *bovenop* het effect in de *nulmeting* voor die treatment. Een sterretje bij de standaardgroep bijvoorbeeld betekent dan een betekenisvolle toe- of afname in de tijd vergeleken met de score in de nulmeting. Bij de beschrijving van de resultaten zal de nadruk liggen op betekenisvolle verschillen in scores die toe te schrijven zijn aan het effect van de treatments. In totaal zijn er 1278 vragenlijsten ingevuld: 507 in meting 1, 514 in meting 2 en 257 in meting 3. Echter, in de tabellen die volgen kan het aantal vragenlijsten van dit totaal afwijken omdat er dan missende antwoorden zijn op bepaalde variabelen in de analyse. De totale aantallen staan in de tabellen vermeld. In totaal hebben 2 verschillende klantmanagers de treatment eigen regie begeleid, 4 klantmanagers intensieve begeleiding, waarvan ook twee de standaardgroep hebben begeleid. In totaal hebben 5 klantmanagers de standaardgroep begeleid.

Motivatie: eigen treatment

In de vragenlijsten zijn de klantmanagers gevraagd om aan te geven in hoeverre zij de deelnemer gemotiveerd achten om zich in te zetten voor respectievelijk de eigen treatment en de (hoogst bereikbare) doelen. De resultaten worden getoond als een rapportcijfer van 0-10, waarbij 0 totaal niet gemotiveerd betekent en 10 heel gemotiveerd.

In tabel 3.2.1 staan de gemiddelde scores op motivatie voor de eigen treatment volgens de klantmanager, per treatmentgroep en meting. Ten eerste kunnen we in de laatste kolom 'gemiddelde totaal' zien dat de treatment eigen regie even hoog scoort als de standaardgroep, maar de intensieve begeleidingsgroep het hoogst scoort. Echter, kijken we naar de significantie, dan zien we alleen een significant effect van de treatment eigen regie, ten opzichte van de standaardgroep. Kijken we naar tabel 1 in Bijlage B dan zien we dat dit een sterk negatief effect is ($b=-3.1$; $p<0,05$).

van het tweede model met interacties zijn te vinden in Bijlage B, tabel 2. Voor het gemak van interpretatie laten we in deze paragraaf de geschatte gemiddelden zien, en niet de volledige uitkomsten van de regressieanalyse. De regressieanalyse (met alle eerdergenoemde predictoren) is wel gebruikt als basis om de gemiddelden in tabel 3.2.1 t/m 3.2.10 te schatten.

¹⁸ Waarom 'deels'? als het effect van een treatment significant is, verklaart dit effect nog niet 100% van de verschillen in gemiddelden tussen groepen. Ook de variabelen klantmanager nummer en zittend bestand verklaren een deel en daarnaast zijn er altijd factoren die een invloed uitoefenen, maar niet worden meegenomen in het model.

Tabel 3.2.1: Gemiddelde scores motivatie eigen treatment (0-10)

	Nulmeting	N	Meting 2	N	Meting 3	N	Gemiddelde Totaal	N totaal
Eigen regie	5,6**	139	5,7***	121	6,1*	58	5,8**	318
Intensieve begeleiding	6,9*	177	6,9***	170	7,4*	98	7,1	445
Standaardgroep	4,9	178	6,4***	185	6,2***	94	5,8	457
Totaal	5,8	494	6,3***	476	6,6**	250	6,2	1220

Noot: ***p<0,01; **p<0,05; *p<0,10, getest op basis van een multiple regressieanalyse (N=1220)

Bron: Vragenlijsten nulmeting en vervolgmetingen procesevaluatie Vertrouwensexperiment Apeldoorn, Tilburg University, 2020.

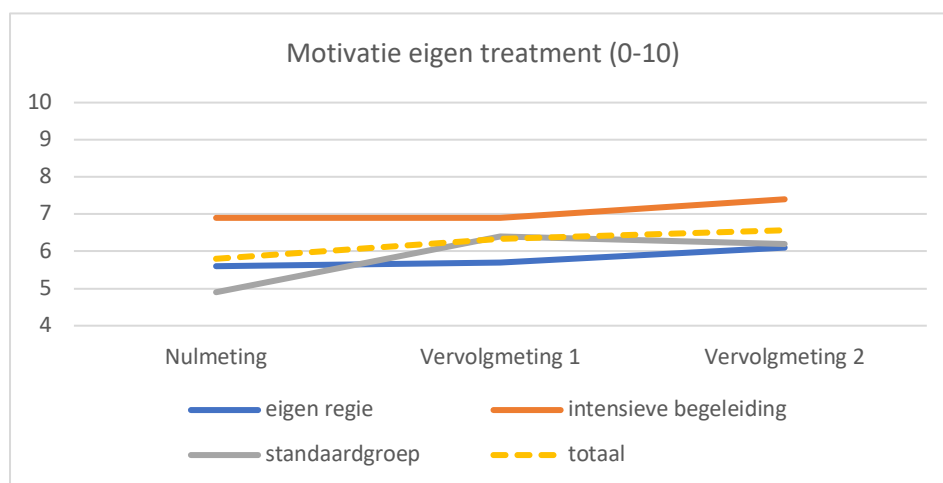
Dit betekent dat in vergelijking met de standaardgroep, de treatment eigen regie op basis van het treatment effect significant slechter scoort op motivatie volgens de klantmanagers. We zien dit niet terug in de gemiddelden, wat kan komen doordat er wellicht andere factoren zijn die een rol spelen in het bepalen van dit gemiddelde, waardoor de verschillen weer kleiner worden. We hebben dit model ook zonder klantmanagernummer getest (hier niet getoond) en dan zien we dat het negatieve effect van eigen regie niet significant is en dat er een significant positief effect is van intensieve begeleiding. Als we dan de modellen vergelijken, dan kunnen we stellen dat het positieve effect van intensieve begeleiding wordt verklaard door welke klantmanager de vragenlijst heeft ingevuld en dat het negatieve effect van de eigen regie treatment pas zichtbaar wordt als we controleren voor welke klantmanager het heeft ingevuld.

Daarnaast hebben we het model met klantmanagernummer ook getest met een groter aantal controlevariabelen¹⁹ en dan blijkt dat er geen effect meer van treatment is, dus dat het negatieve effect van eigen regie wordt verklaard door compositieverschillen tussen de treatmentgroepen. Met name de duur van de uitkering heeft een negatief effect. Dit is echter wel op een kleinere steekproef getest (N=995), vanwege missende waarden. Het effect van eigen regie is dus niet robuust.

Ten tweede zien we over alle groepen heen een stijging van de motivatie over tijd van 5,8 naar 6,6. Deze stijging is significant en positief (zie ook tabel 1, Bijlage B). Als we dan naar de tussenliggende cellen kijken, dan zien we een stijging in de gemiddelde motivatie voor alle groepen over tijd, maar nader bekeken blijkt er uit tabel 2 in Bijlage B dat alleen het effect van treatment voor de standaardgroep positief is over tijd, en voor de eigen regiegroep sterk negatief (in vergelijking met het verschil met de standaardgroep in de nulmeting). Het effect voor de intensieve begeleidingsgroep begint positief (in vergelijking met de standaardgroep), maar wordt later in het experiment negatief. De stijging in de gemiddelde voor de groepen eigen regie en intensieve begeleiding wordt dus door iets anders veroorzaakt dan door het treatment effect. Opvallend is dat het effect van de standaardgroep positief is, terwijl de andere twee treatments juist negatief uitwerken op de motivatie. Wel moet bij de resultaten voor de derde meting rekening gehouden worden met dat op het moment van analyse slechts ongeveer de helft van de vragenlijsten is ingevuld en dat dit waarschijnlijk een selectieve groep is, die na twee jaar nog steeds in

¹⁹ Leeftijd, geslacht, opleidingsniveau, huisvesting, leefvorm, duur van uitkering, nationaliteit

het experiment zit²⁰. Klantmanagers vullen voor uitgestroomde deelnemers in beginsel geen vragenlijsten meer in nadat ze zijn uitgestroomd. Ook hier hebben we een model gedraaid zonder klantmanagernummer en dan blijkt weer dat de treatments eigen regie en intensieve begeleiding een positief effect hebben ten opzichte van de standaardgroep in de nulmeting, maar dit verandert in een negatief effect als we de invloed van de klantmanager eruit filteren. De resultaten voor het treatment effect zijn dus sterk afhankelijk van de klantmanager die de vragenlijst invult. In figuur 3.2.1 zijn de gemiddelde voor de groepen over tijd weergegeven.



Figuur 3.2.1. Gemiddelde score op motivatie voor eigen treatment per treatmentgroep en meting

Motivatie: hoogst bereikbare doel

Daarnaast hebben de betrokken klantmanagers ook ingevuld wat hun inschatting is van de motivatie van de deelnemer om zich in te zetten voor de (hoogst bereikbare) doelen. In tabel 3.2.2 zien we dat de treatments eigen regie en intensieve begeleiding iets hoger scoren dan de standaardgroep (kolom gemiddelde totaal). Echter alleen bij de eigen regie groep is er een significant verschil met de standaardgroep en het effect is negatief (zie tabel 1, Bijlage B). Dat de groep eigen regie hoger scoort dan de standaardgroep, is dus niet toe te schrijven aan het effect van de treatment, omdat dat juist negatief is. Ook hier hebben we weer een model getest zonder het effect van klantmanager en dan zien we in dat model dat de effecten van eigen regie en intensief significant en positief zijn, maar negatief worden als we corrigeren voor klantmanager. Ook hier spelen de verschillen tussen de klantmanagers dus weer een belangrijke rol in het effect van de treatments. Ook hebben we een model getest met de uitgebreide set controlevariabelen en dan vervalt het effect van eigen regie. Deze is dus niet robuust.

²⁰ De klantmanagers waren op het moment van analyse nog bezig met de verwerking van een aantal procesevaluatievragenlijsten. Deze aanvullende gegevens zullen in de definitieve rapportage worden opgenomen.

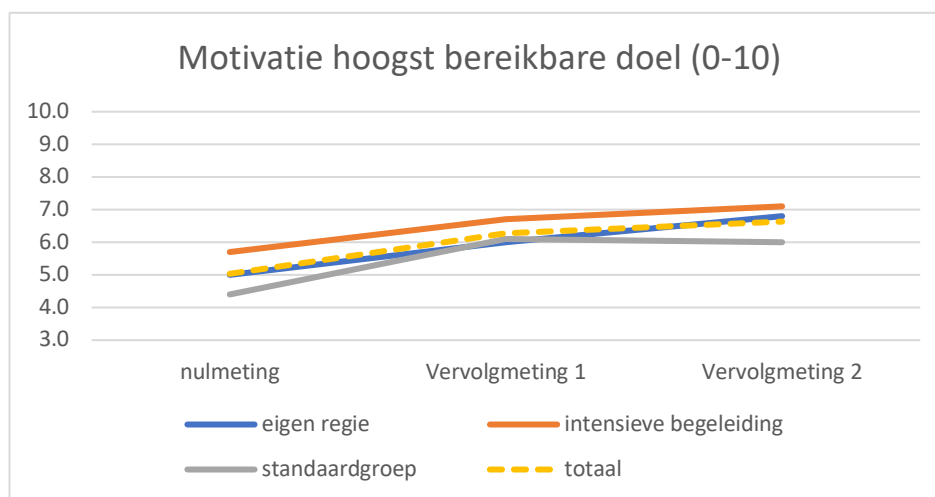
Tabel 3.2.2: Gemiddelde scores op motivatie hoogst bereikbare doel (0-10)

	Nulmeting	N	Meting 2	N	Meting 3	N	Gemiddelde	N
Eigen regie	5,0	131	6,0*	96	6,8	58	5,9*	285
Intensieve begeleiding	5,7	175	6,7*	165	7,1	99	6,5	439
Standaardgroep	4,4	173	6,1***	181	6,0***	94	5,5	448
Totaal	5,0	479	6,3***	442	6,6***	251	6,0	1172

Noot: ***p<0,01; **p<0,05; *p<0,10, getest op basis van een multiple regressieanalyse (N=1172)

Bron: Vragenlijsten nulmeting en vervolgmetingen procesevaluatie Vertrouwensexperiment Apeldoorn, Tilburg University, 2020.

Wat wel robuust is zijn de positieve effecten over tijd. Over de groepen heen vindt er een stijging plaats in motivatie en deze is substantieel en significant (zie tabel 1, Bijlage B). Ook blijft deze overeind in de modellen met extra controlevariabelen. Kijken we naar de tussenliggende cellen, dan zien we dat voor de standaardgroep er een significante stijging in motivatie is in de tweede en derde meting ten opzichte van de nulmeting. En voor de intensieve en eigen regie groepen zien we ook een significant effect, maar deze blijkt negatief te zijn (tabel 2, Bijlage B). De hogere score die we dus in tabel 3.2.2 zien voor de twee treatments wordt dus door iets anders veroorzaakt, dan door het treatment effect.



Figuur 3.2.2. Gemiddelde score op motivatie voor het hoogst bereikbare doel per treatmentgroep en meting

Realistische eisen

Het eerste aspect van de begeleiding waar we nu aandacht aan besteden is hoe realistisch de klantmanagers de eisen vinden die de deelnemer aan de te behalen doelen stelt. Ook hier kijken we weer naar een rapportcijfer van 0-10 waarin 0 staat voor zeer onrealistisch en 10 voor zeer realistisch.

In tabel 3.2.3 zien we de resultaten voor de drie treatmentgroepen. In de laatste kolom zien we dat de eigen regie en intensieve begeleidingsgroep iets hoger scoren dan de standaardgroep, maar deze verschillen kunnen niet aan de treatments worden toegerekend. Ook hier hebben we weer een test gedaan zonder te controleren voor klantmanager en dan blijken de effecten van eigen regie en intensieve begeleiding wel significant en positief. Maar als we klantmanager wel meenemen dan verdwijnen deze effecten en moeten we de verschillen tussen de groepen dus toeschrijven aan het effect van de klantmanager. In het model met de uitgebreide set controlevariabelen vinden we ook geen verschillen tussen de treatments.

Tabel 3.2.3: Gemiddelde scores op de mate van realistisch eisen stellen (0-10)

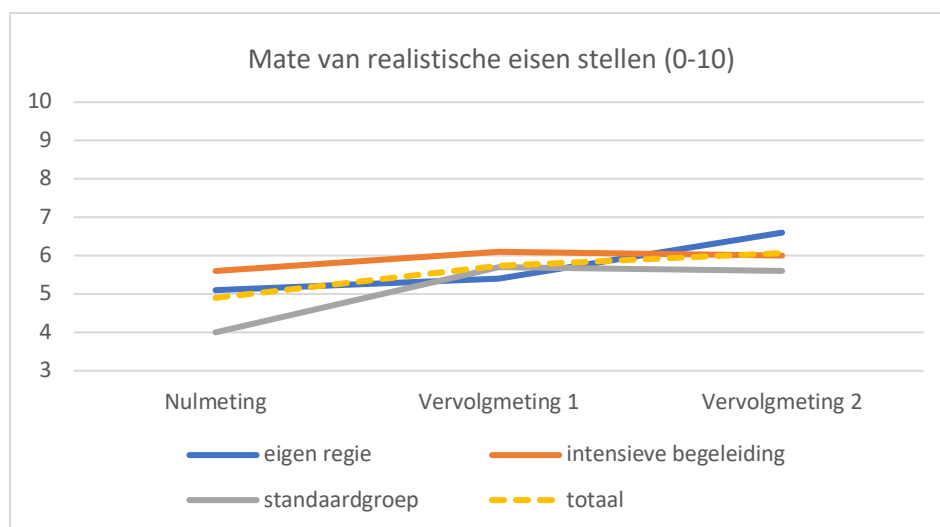
	Nulmeting	N	Meting 2	N	Meting 3	N	Gemidd elde	N
Eigen regie	5,1	130	5,4***	103	6,6	88	5,7	321
Intensieve begeleiding	5,6**	174	6,1**	148	6**	88	5,9	410
Standaardgroep	4	174	5,7***	185	5,6***	92	5,1	451
Totaal	4,9	478	5,7***	436	6,1***	268	5,6	1182

Noot: ***p<0,01; **p<0,05; *p<0,10, getest op basis van een multiple regressieanalyse (N=1182)

Bron: Vragenlijsten nulmeting en vervolgmetingen procesevaluatie Vertrouwensexperiment Apeldoorn, Tilburg University, 2020.

Verder zien we wederom over alle groepen heen een significante positieve stijging in de mate van realistisch eisen stellen (zie ook tabel 1, Bijlage B). Ook in het uitgebreide model met meer controlevariabelen blijft dit effect van tijd overeind. Als we dan kijken naar wat er verandert in de tijd voor welke groep, zien we voor de standaardgroep dat er weer een significant treatment effect is dat zich uit in een positieve toename over tijd (ten opzichte van de nulmeting). Voor eigen regie en intensieve begeleiding zien we ook een stijging, maar tabel 2 in Bijlage B laat zien dat het treatment effect zelf hier weer negatief is. Intensieve begeleiding scoort weliswaar significant hoger dan de standaardgroep in de nulmeting (ook in het model met meer controlevariabelen), maar in de latere metingen wordt dit positieve effect tenietgedaan door een negatief interactie-effect.

De ontwikkelingen uit tabel 3.2.3 worden ook getoond in figuur 3.2.3.



Figuur 3.2.3. Gemiddelde scores op de mate van realistisch eisen stellen per treatmentgroep en meting

Zelf in actie komen

Een ander aspect van de begeleiding is de mate waarin de klantmanager de deelnemer in staat acht om zelf in actie te komen om de gestelde doelen te behalen. In onderstaande tabel zien we de resultaten. Hier is weer een schaal van 0-10 gebruikt, waarbij 0 staat voor zeer slecht en 10 voor zeer goed.

In tabel 3.2.4 zien we in de laatste kolom dat de eigen regie groep een lager gemiddelde heeft dan de standaardgroep, terwijl intensieve begeleiding iets hoger scoort. Uit tabel 1 uit bijlage B blijkt dat eigen regie significant lager scoort op het zelf in actie komen dan de standaardgroep als gevolg van het effect van de treatment. Over de tijd zien we echter wel weer een significante toename over alle groepen heen, van 3,9 naar 5,6.

Tabel 3.2.4: Gemiddelde scores op zelf in actie komen (0-10)

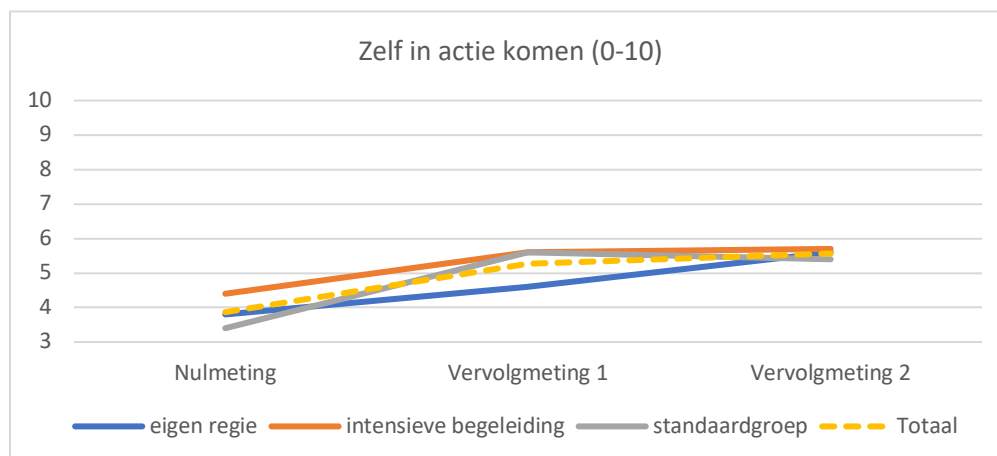
	Nulmeting	N	Meting 2		Meting 3		Gemiddelde	N
Eigen regie	3,8	127	4,6**	129	5,6	60	4,7**	316
Intensieve begeleiding	4,4	159	5,6**	165	5,7	99	5,2	423
Standaardgroep	3,4	176	5,6***	186	5,4***	94	4,8	456
Totaal	3,9	462	5,3***	480	5,6***	253	4,9	1195

Noot: ***p<0,01; **p<0,05; *p<0,10, getest op basis van een multiple regressieanalyse (N=1195)

Bron: Vragenlijsten nulmeting en vervolgmetingen procesevaluatie Vertrouwensexperiment Apeldoorn, Tilburg University, 2020.

Nader bekeken blijkt dat deze positieve toename vooral toe te schrijven is aan het positieve effect van de standaardtreatment (zie ook tabel 2, Bijlage B). Indien de standaard treatment niet verandert over tijd zouden we deze stijging niet verwachten. De significante effecten voor eigen regie en intensieve begeleiding in meting 2 blijken negatief te zijn ten opzichte van de nulmeting, terwijl we wel een stijging

zien in de scores op zelf in actie komen. Deze stijging heeft dus een andere oorzaak. Zonder controle voor klantmanager/consulent vinden we een positief en significant effect voor intensieve begeleiding en geen effect voor eigen regie. Wederom heeft de klantmanager een belangrijk effect op de gemiddelde score die wordt gerapporteerd en wordt het treatment effect van eigen regie negatief en verdwijnt het effect van intensieve begeleiding als we wel controleren voor klantmanager. In het uitgebreide model met controlevariabelen verdwijnt ook het negatieve effect van eigen regie (in de laatste kolom), maar blijven de negatieve interactie-effecten in meting 2 voor beide treatments wel overeind.



Figuur 3.2.4. Gemiddelde scores op zelf in actie komen per treatmentgroep en meting

In figuur 3.2.4 zijn deze resultaten ook te zien en wordt duidelijk dat alle groepen deelnemers volgens de klantmanagers een onvoldoende scoren op het zelf in actie komen, maar dat er wel sprake is van een stijgende lijn.

Zelf doelen stellen

Een derde aspect van de begeleiding is in hoeverre deelnemers zelf in staat zijn doelen te stellen, volgens de klantmanager/consulent. Ook hier is weer een score van 0 tot 10 gebruikt, waarbij 0 betekent geheel niet in staat zelf doelen te stellen en 10 geheel wel. In tabel 3.2.5 staan de resultaten voor de treatmentgroepen over de tijd. We zien kleine verschillen tussen de groepen in de laatste kolom, maar deze zijn niet toe te schrijven aan het effect van de treatments.

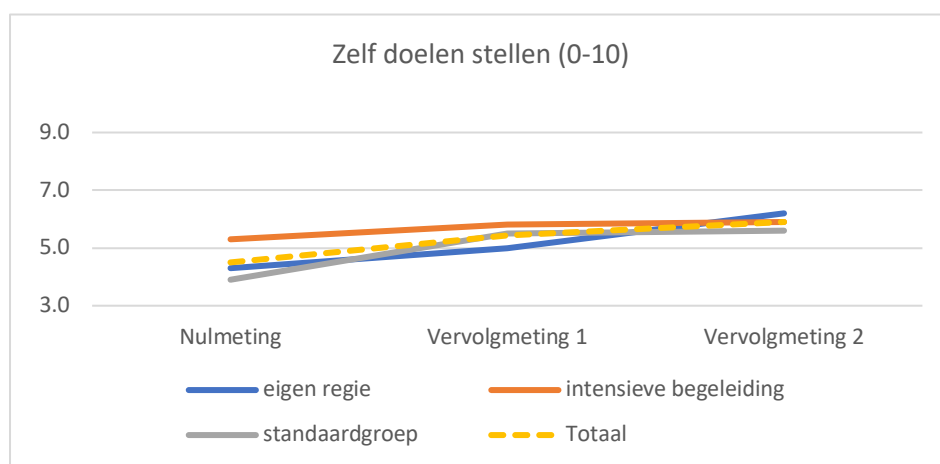
Tabel 3.2.5: Gemiddelde scores op zelf doelen stellen (0-10)

	Nulmeting	N	Meting 2	N	Meting 3	N	Gemiddelde	N
Eigen regie	4,3	135	5,0**	122	6,2	60	5,2	317
Intensieve begeleiding	5,3**	164	5,8**	163	5,9**	98	5,7	425
Standaardgroep	3,9	178	5,5***	188	5,6***	94	5,0	460
Totaal	4,5	477	5,4***	473	5,9***	252	5,3	1202

Noot: *** $p < 0,01$; ** $p < 0,05$; * $p < 0,10$, getest op basis van een multiple regressieanalyse (N=1202)

Bron: Vragenlijsten nulmeting en vervolgmetingen procesevaluatie Vertrouwensexperiment Apeldoorn, Tilburg University, 2020.

Als we niet controleren voor klantmanager (model hier niet getoond), dan zien we een sterk significant en positief effect voor intensieve begeleiding, maar deze wordt geheel verklaard door het effect van klantmanager in het hier getoonde model. Wederom zien we een positief en significant stijging over de tijd, die ook overeind blijft als we het model testen met de uitgebreide set controlevariabelen. In de tussenliggende cellen vinden we voor de standaardgroep weer een positief en significant effect van treatment over tijd. Ook blijkt dat de intensieve begeleidingsgroep in de nulmeting significant hoger scoort dan de standaardgroep, maar dat het treatment effect daarna negatief wordt (zie tabel 2, Bijlage B). Het treatment effect van eigen regie is ook negatief in meting 2, ondanks dat de gemiddelde score wel stijgt, maar onder het gemiddelde blijft van de standaardgroep. Ook in figuur 3.2.5 zien we een stijging van met name de eigen regie groep, maar juist deze groep kent een negatief significant effect van treatment in meting twee (zie tabel 2, Bijlage B), de stijging heeft dus een andere oorzaak. De standaardgroep stijgt dus in dezelfde mate mee als de andere groepen terwijl een stabiel patroon verwacht had mogen worden indien de treatment hetzelfde zou zijn gebleven.



Figuur 3.2.5. Gemiddelde scores op zelf doelen stellen per treatmentgroep en meting

Verwacht effect van treatment op cliënt

Ook is de klantmanagers gevraagd hoe groot zij verwachten dat het effect van de treatment is voor de deelnemer op een schaal van 0 tot 10. Hierbij staat 0 voor totaal geen effect en 10 voor een zeer groot effect. In tabel 3.2.6 zien we opvallend genoeg dat klantmanagers de verwachte treatment effecten bij de eigen regie en intensieve begeleidingsgroep lager scoren dan bij de standaardgroep. Voor de eigen regie groep vinden we dan ook een sterk negatief en significant effect in vergelijking met de standaardgroep.

Dit is opvallend aangezien het experiment juist is opgezet met de verwachting dat de treatments eigen regie en intensieve begeleiding betere resultaten zouden opleveren voor de deelnemers. Ook zien we voor het eerst een daling van de scores over tijd en deze zijn significant (zie ook tabel 1, Bijlage B). Dit suggereert dat de consultants in Apeldoorn kennelijk gehecht zijn aan de standaardmethoden die zij gewend zijn te gebruiken. Dat betekent voor het onderzoek dat door de habitus en opstelling van de

consulenten de theoretisch ontworpen treatments mogelijk minder zuiver zijn uitgevoerd dan met het experiment is beoogd. Dit zou deels kunnen verklaren waarom de verschillen met de standaard treatment op de andere aspecten veelal niet significant zijn.

Tabel 3.2.6: Gemiddelde scores op het verwachte effect van de treatment (0-10)

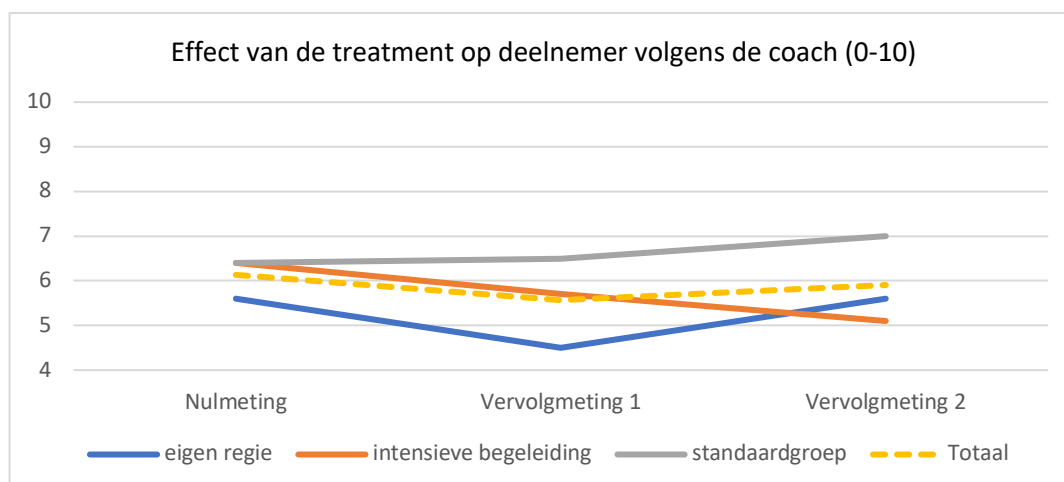
	Nulmeting	N	Meting 2	N	Meting 3	N	Gemiddelde	N
Eigen regie	5,6**	143	4,5***	133	5,6	61	5,2**	337
Intensieve begeleiding	6,4	180	5,7**	162	5,1***	96	5,7	438
Standaardgroep	6,4	182	6,5	188	7*	93	6,6	463
Totaal	6,1	505	5,6***	483	5,9*	250	5,9	1238

Noot: ***p<0,01; **p<0,05; *p<0,10, getest op basis van een multiple regressieanalyse (N=1238)

Bron: Vragenlijsten nulmeting en vervolgmetingen procesevaluatie Vertrouwensexperiment Apeldoorn, Tilburg University, 2020.

Als we dan naar de tussenliggende cellen kijken, dan zien we dat de eigen regie groep in de nulmeting al lager scoort dan de standaardgroep en dat dit (deels) toe te schrijven is aan een negatief treatment effect (zie tabel 2, Bijlage B). De aanvangsverwachting van de klantmanagers was dus al bij de start (na zo'n 2 tot 3 maanden) lager dan die voor de andere twee treatments. Vervolgens zien we voor de standaardgroep nog een licht positief effect in de derde meting, maar voor de andere twee treatments vinden we alleen nog negatieve effecten.

Indien we een model schatten zonder klantmanager, dan vinden we negatieve effecten voor zowel eigen regie als intensief, maar die van intensieve begeleiding verdwijnt als we klantmanager toevoegen aan het model. Als we controleren voor de grotere set controlevariabelen, dan vinden we geen effecten meer van de treatments (de laatste kolom), maar voor de tussenliggende cellen wel negatieve effecten van eigen regie in meting 2 en voor intensief in meting 2 en 3. Al met al zijn de verwachtingen voor de effecten van de treatments laag en nemen deze over de tijd verder af, behalve voor de standaardgroep, die toch al het hoogste scoort. Dit zou kunnen betekenen dat in de loop van de tijd de treatments juist meer op elkaar zijn gaan lijken waardoor een zuivere vergelijking niet mogelijk is.



Figuur 3.2.6. Gemiddelde scores op het verwachte effect van de treatment per treatmentgroep en meting

Sociale participatie

Wat betreft sociale participatie, zijn klantmanagers gevraagd om een inschatting te geven van de mate hiervan bij de deelnemer. Deze vraag kent weer scores van 0 tot 10, waar 0 staat voor participeert totaal niet en 10 staat voor participeert totaal wel in de perceptie althans van de klantmanager.

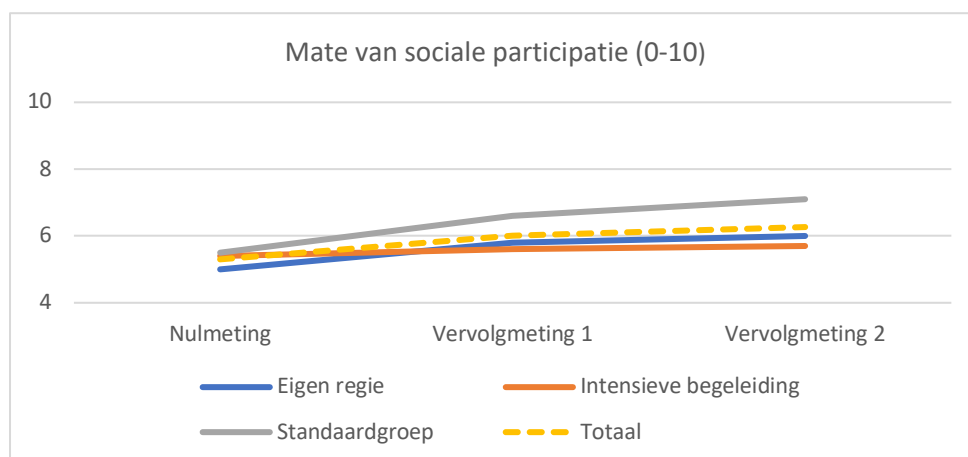
Tabel 3.2.7: Gemiddelde scores op sociale participatie (0-10)

	Nulmeting	N	Meting 2	N	Meting 3	N	Gemiddelde	N
Eigen regie	5*	143	5,8	124	6	57	5,6**	324
Intensieve begeleiding	5,4	180	5,6**	165	5,7**	89	5,6	434
Standaardgroep	5,5	182	6,6***	185	7,1***	89	6,4	456
Totaal	5,3	505	6,0***	474	6,3***	235	5,9	1214

Noot: *** $p < 0,01$; ** $p < 0,05$; * $p < 0,10$, getest op basis van een multiple regressieanalyse ($N = 1214$)

Bron: Vragenlijsten nulmeting en vervolgmetingen procesevaluatie Vertrouwensexperiment Apeldoorn, Tilburg University, 2020.

In tabel 3.2.7 zien we dat voor de groepen eigen regie en intensieve begeleiding de mate van sociale participatie lager wordt ingeschat dan voor de standaardgroep (kolom gemiddelde totaal). Dit is ook duidelijk te zien in figuur 3.2.7. In tabel 1 (Bijlage B) is te zien dat de treatment eigen regie een significant negatief effect heeft op de mate van sociale participatie vergeleken met de standaardgroep. Wel zien we hier weer een positief effect over de tijd voor alle groepen (onderste rij). Er vindt over alle groepen heen een significante stijging plaats in de mate van sociale participatie. In het model waarin we niet controleren voor klantmanager (hier niet getoond), vinden we een sterk negatief en significant effect voor de treatments eigen regie en intensieve begeleiding. Echter, het effect van intensieve begeleiding verdwijnt dus als we klantmanager meenemen in het model en wordt dus verklaard door het effect van de klantmanager op de scores. Controleren we ook nog voor de ruimere set controlevariabelen, dan vinden we geen effecten meer voor de treatments, maar wel voor de verandering over tijd.



Figuur 3.2.7. *Gemiddelde scores op sociale participatie per treatmentgroep en meting*

In de tussenliggende cellen vinden we wederom een negatief effect van eigen regie in de nulmeting. Wellicht is hier sprake van een selectie-effect, omdat er tijdens de nulmeting nog niet veel begeleiding had plaatsgevonden. Daarnaast zien we een significante stijging voor de standaardgroep in de tweede en derde meting, ten opzichte van de nulmeting. Zij gaan in de ogen van de klantmanager dus meer participeren. Verder zien we voor de intensieve begeleidingsgroep dat er een stijging plaatsvindt in de tweede en derde meting, maar dat beide treatment effecten negatief zijn (zie tabel 2, Bijlage B). De treatment zelf werkt dus niet bevorderend voor de sociale participatie. Deze negatieve effecten voor intensieve begeleiding blijven bestaan ook als we controleren voor de uitgebreide set controlevariabelen.

Participatieladder

Naast de sociale participatie is de klantmanagers ook gevraagd een score te geven voor de afstand van de deelnemer tot de arbeidsmarkt op een schaal van 1 tot 6. Dit wordt ook wel de participatieladder of tredeladder genoemd. 1 staat hier voor een zeer grote afstand tot de arbeidsmarkt, terwijl 6 staat voor het verrichten van betaald werk. De participatieladder speelt een belangrijke rol in het begeleidingsproces aangezien hiermede de vooruitgang in de ondersteuning kan worden gemeten. De participatieladder vormt onderdeel van de registratie van de procesbewaking en beheer en is als zodanig ook opgenomen in de BUS (bijstandsuitkeringsstatistiek) database en statistiek. De consultants worden geacht deze registratie zorgvuldig op basis van nauwkeurig omschreven criteria bij te houden. Ze doen dat op gezette tijden. In het onderzoek is aan hen gevraagd dit te doen op elk van de drie momenten dat ze een vragenlijst hebben ingevuld waarin hen gevraagd is naar het proces in de afgelopen periode. In deze zin geeft het een zo objectief mogelijk beeld van de vorderingen of niet van de bijstandsgerechtigde in het proces van begeleiding. Voor de gemeente is het de peilstok voor een meer of minder succesvol re-integratie -en begeleidingsproces van de betrokkene.

Omdat het hier een schaal van 1 tot 6 betreft, moet er rekening mee worden gehouden dat een score hoger dan drie 'voldoende' is. Het gemiddelde over alle groepen ligt op 3,6, dus net voldoende maar wel met een zekere afstand tot de arbeidsmarkt

Tabel 3.2.8: Gemiddelde positie op participatieladder (1-6)

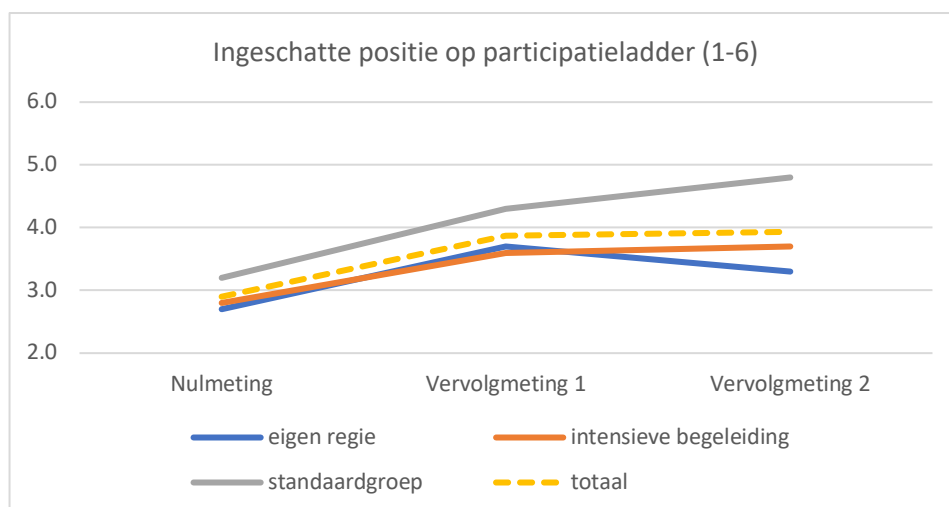
	Nulmeting	N	Meting 2	N	Meting 3	N	Gemiddelde totaal	N
Eigen regie	2,7**	141	3,7	134	3,3***	61	3,2**	336
Intensieve begeleiding	2,8	179	3,6**	172	3,7**	99	3,4	450
Standaardgroep	3,2	178	4,3***	187	4,8***	93	4,1	458
Totaal	2,9	498	3,9***	493	3,9***	253	3,6	1244

Noot: *** $p < 0,01$; ** $p < 0,05$; * $p < 0,10$, getest op basis van een multiple regressieanalyse (N=1244)

Bron: Vragenlijsten nulmeting en vervolgmetingen procesevaluatie Vertrouwensexperiment Apeldoorn, Tilburg University, 2020.

De standaardgroep scoort opnieuw het hoogste met een 4,1 en de eigen regie groep het laagste met een 3,2. Voor de eigen regie groep vinden we dan ook een significant negatief treatment effect in vergelijking met de standaardgroep (tabel 1, Bijlage B). In het model zonder de controle voor klantmanager wat we hebben getest (hier niet getoond) vinden we sterk significante negatieve effecten voor zowel eigen regie en intensief, maar voor intensieve begeleiding wordt dit effect dus weg verklaard door klantmanager in ons huidige model. Weer speelt de klantmanager dus een belangrijke rol. Controleren we voor de uitgebreide set controlevariabelen, dan valt ook het negatieve treatment effect van eigen regie weg en dat wordt dus verklaard door verschillen in compositie tussen de groepen.

We vinden ook een significante stijging over tijd over alle groepen heen (onderste rij), ten opzichte van de nulmeting, die ook overeind blijven in het model met extra controlevariabelen. Deze stijging is weer toe te rekenen aan de significante stijging die de standaardgroep doormaakt, mede als gevolg van de treatment. In de nulmeting blijkt de eigen regie groep al significant lager te scoren dan de standaardgroep, terwijl er nog nauwelijks begeleiding had plaatsgevonden. Dit effect valt dan ook weg als we voor de uitgebreide set controlevariabelen controleren en kan dus worden toegeschreven aan compositieverschillen tussen de groepen. Ondanks dat de score van de intensieve begeleidingsgroep stijgt ten opzichte van de nulmeting, is het treatment effect wel negatief, wat wil zeggen dat ten gevolge van de treatment, de positie op de participatieladder afneemt voor de intensieve begeleidingsgroep. Hetzelfde geldt voor de eigen regie groep in de derde meting. Figuur 3.2.8 geeft duidelijk de positie van de standaardgroep weer ten opzichte van de andere twee groepen.



Figuur 3.2.8. Gemiddelde positie op participatieladder per treatmentgroep en meting

Leereffecten

Het begeleiden van de deelnemers met behulp van de treatments is een leerproces voor zowel de deelnemer als voor de klantmanager. Om dit leerproces in kaart te brengen gedurende het experiment, is aan de klantmanagers de vraag gesteld in hoeverre zij vinden dat de deelnemer leert om zelfredzaam te zijn en in hoeverre zij zelf leren van de treatment om zo de deelnemer beter te begeleiden. Beide vragen hebben een score van 0 tot 10 gekregen, waarbij voor het leereffect van de deelnemer 0 staat voor totaal niet zelfredzaam en 10 voor volledig zelfredzaam en voor het leereffect van de klantmanager staat 0 voor niets bijgeleerd en 10 voor heel veel bijgeleerd.

Leereffecten klant: zelfredzaamheid

In tabel 3.2.9 staan de resultaten voor de leereffecten van de deelnemers. Opvallend is dat we kleine verschillen zien tussen de treatmentgroepen, waarbij de standaardgroep het hoogste scoort, maar dat we deze niet kunnen verklaren aan de hand van de effecten van de treatments zelf.

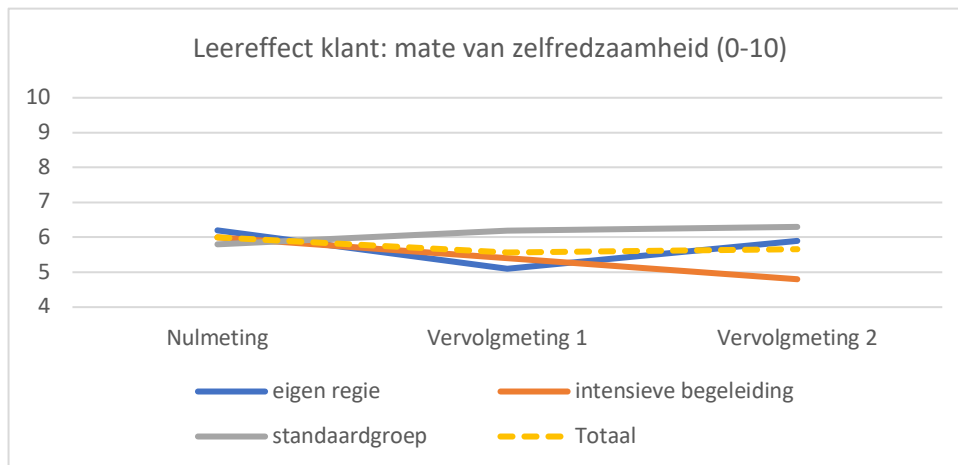
Tabel 3.2.9: Gemiddelde ingeschatte mate van zelfredzaamheid (leereffect klant) (0-10)

	Nulmeting	N	Meting 2	N	Meting 3	N	Gemiddelde	N
Eigen regie	6,2	143	5,1***	120	5,9	59	5,7	322
Intensieve begeleiding	6	180	5,4**	164	4,8***	96	5,4	440
Standaardgroep	5,8	182	6,2*	189	6,3	93	6,1	464
Totaal	6	505	5,6**	473	5,7**	248	5,7	1226

Noot: ***p<0,01; **p<0,05; *p<0,10, getest op basis van een multiple regressieanalyse (N=1226)

Bron: Vragenlijsten nulmeting en vervolgmetingen procesevaluatie Vertrouwensexperiment Apeldoorn, Tilburg University, 2020.

In het model waarin we niet voor klantmanager hebben gecontroleerd (hier niet getoond) vinden we wel significante negatieve effecten van beide treatments ten opzichte van de standaardgroep, maar die worden dus in dit huidige model verklaard door het effect van de klantmanager (waarbij de effecten voor sommige klantmanagers significant zijn). Daarnaast zien we een afnemende mate van gepercipieerde zelfredzaamheid over tijd en deze daling is significant. Als we naar de tussenliggende cellen kijken, dan zien we dat er voor de standaardgroep een lichte stijging plaatsvindt in zelfredzaamheid over de tijd, terwijl we voor de andere twee groepen alleen een afname zien over de tijd, die ook significant is toe te schrijven aan het effect van de treatment (zie ook tabel 2, Bijlage B).



Figuur 3.2.9. Gemiddelde ingeschatte mate van zelfredzaamheid (leereffect klant) per treatmentgroep en meting

In Figuur 3.2.9 is te zien dat aanvankelijk de verschillen tussen de groepen klein zijn, maar naarmate het experiment vordert, de scores voor de standaardgroep stijgen en hoger worden dan bij de andere twee groepen. Voornamelijk de scores van de eigen regie groep daalt flink. Echter, hierbij moet weer worden bedacht dat de derde meting voorsnog slechts voor de helft (waarschijnlijk de minder kansrijke) deelnemers is ingevuld. Uit deze bevindingen leiden we opnieuw af dat de standaard treatment mee is geëvolueerd met de andere treatments waardoor een zuivere vergelijking is bemoeilijkt.

Leereffecten klantmanager

Vervolgens kijken we naar het leereffect voor de klantmanager/consulent. Ook deze keer vinden we geen verschil tussen de treatmentgroepen die toe te schrijven is aan het effect van de treatments (laatste kolom). Wel zien we dat de eigen regie groep het hoogste scoort en daar leert de klantmanager dus blijkbaar het meeste van. Verder zien we wel een significante stijging van de inschatting van het leereffect over de tijd (zie ook tabel 1, Bijlage B). Deze stijging is toe te schrijven aan de significante stijging die de standaardgroep doormaakt als gevolg van de treatment in meting 2 en meting 3. Verder blijkt dat het effect van eigen regie in de nulmeting negatief is (zie tabel 2, Bijlage B), ondanks dat de score iets hoger is dan in de standaardgroep. We zien ook een stijging voor intensieve begeleiding in de tweede en derde meting, maar dat wordt niet veroorzaakt door de treatment, want dat effect is significant negatief.

Daarnaast zien we eerst een stijging en dan een daling voor de eigen regie groep (zie ook figuur 3.2.10), maar het effect van eigen regie is in zowel de tweede als derde meting (sterk) negatief.

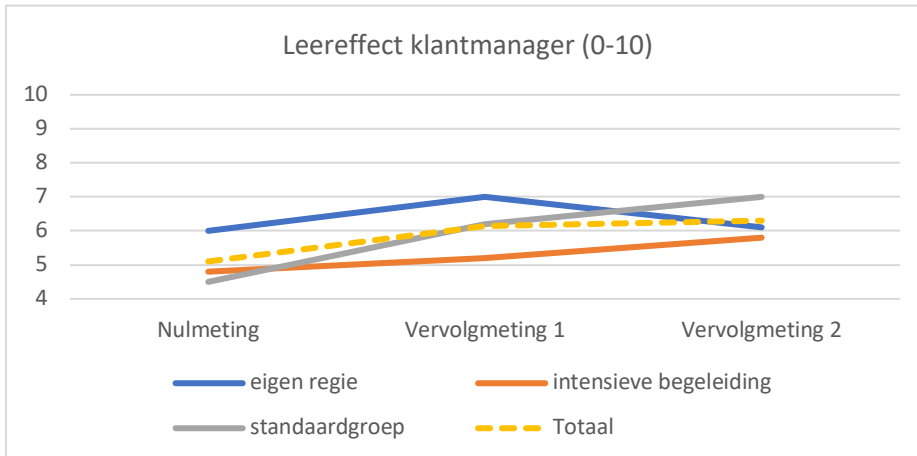
Tabel 3.2.10: Gemiddelde score op leereffect van klantmanager (0-10)

	Nulmeting	N	Meting 2	N	Meting 3	N	Gemiddelde	N
Eigen regie	6	143	7**	142	6,1***	62	6,4	347
Intensieve begeleiding	4,8**	180	5,2***	176	5,8**	100	5,3	456
Standaardgroep	4,5	182	6,2***	195	7***	94	5,9	471
Totaal	5,1	505	6,1***	513	6,3***	256	5,8	1274

Noot: *** $p < 0,01$; ** $p < 0,05$; * $p < 0,10$, getest op basis van een multiple regressieanalyse (N=1274)

Bron: Vragenlijsten nulmeting en vervolgmetingen procesevaluatie Vertrouwensexperiment Apeldoorn, Tilburg University, 2020.

Als we niet zouden controleren voor klantmanager, dan vinden we een significant positief treatment effect voor eigen regie en een significant negatief treatment effect voor intensieve begeleiding, in de kolom 'gemiddelde totaal'. Maar deze treatment effecten worden dus verklaard door het effect van de klantmanagers in het huidige model. Het is dus geen puur treatment effect, maar een effect veroorzaakt door de persoon van de klantmanager die de vragenlijst invult.



Figuur 3.2.10. Gemiddelde score op leereffect van klantmanager per treatmentgroep en meting

Vergelijking eigen regie en intensieve begeleidingsgroep

In voorgaande tabellen is niet beschreven of de verschillen tussen de groepen eigen regie en intensieve begeleiding betekenisvol konden worden toegeschreven aan het effect van treatment, omdat we steeds een vergelijking hebben gemaakt tussen deze twee groepen en de standaardgroep. Dat was onze primaire vergelijking. Echter, om meer inzicht te bieden in het verschil tussen de eigen regie en de intensieve begeleidingsgroep, hebben we een kort overzicht gemaakt in onderstaande tabel. Daarin is steeds de

groep intensieve begeleiding vergeleken met de eigen regie groep en of eventuele verschillen (deels) kunnen worden toegeschreven aan verschillen in treatment. Het gaat hier om de kolom 'gemiddelde totaal' in de tabellen 3.2.1 t/m 3.2.10.

Tabel 3.2.11 Geteste verschillen in score op concepten tussen intensieve begeleiding en eigen regie

Concept	Vershil intensief vs. eigen regie
Motivatie eigen treatment	**
Motivatie hoogst bereikbare doel	n.C.
Realistische eisen	n.s.
Zelf in actie komen	n.s.
Zelf doelen stellen	n.s.
Verwacht effect treatment	**
Sociale participatie	**
Participatieladder	**
Leereffect klant	n.s.
Leereffect klantmanager	n.s.

Noot: *** $p < 0,01$; ** $p < 0,05$; * $p < 0,10$; n.s. (niet significant)

Bron: Vragenlijsten nulmeting en vervolgmetingen procesevaluatie Vertrouwensexperiment Apeldoorn, Tilburg University, 2020.

Er blijkt uit tabel 3.2.11 dat in een aantal gevallen de verschillen tussen deze twee treatmentgroepen significant konden worden toegeschreven aan de treatments. Zo blijkt er een significant verschil te zijn in de motivatie voor de eigen treatment: intensieve begeleiding scoort hoger. Ook voor het verwachte effect van de treatment, de mate van sociale participatie en de participatieladder is er een significant verschil tussen de groepen die toe te schrijven is aan het treatment effect. Op al deze uitkomstmaten is er een positief effect van intensieve begeleiding ten opzichte van de eigen regie groep gevonden.

Vergelijking Apeldoorn en Epe

In alle regressieanalyses modellen die we hebben getest en die hiervoor besproken zijn, hebben we ook gecontroleerd voor uit welke gemeente een deelnemer afkomstig is: Apeldoorn of Epe. Hierdoor hebben we ook zicht op of er verschillen zijn tussen de gemeenten op de besproken uitkomstmaten. Deze zullen hier kort worden besproken. Uit de analyses is gebleken dat deelnemers uit de gemeente Epe op basis van de ingevulde vragenlijsten door de klantmanagers, significant lager scoren op de motivatie voor de eigen treatment, de mate van realistisch eisen stellen en het zelf doelen stellen, dan Apeldoorn. Deze resultaten bestaan, ondanks dat er gecontroleerd is voor klantmanager en ze blijven ook overeind als we de uitgebreide set van controlevariabelen gebruiken. De lagere score voor deelnemers uit Epe op deze items kan dus niet worden toegeschreven aan verschillen tussen klantmanagers of verschillen in compositie van de deelnemers van Apeldoorn en Epe. Deelnemers uit Epe blijken dus, in de ogen van de klantmanager, over het algemeen minder gemotiveerd voor de treatment, minder goed in het stellen van realistische eisen en in het zelf doelen stellen.

Geschiktheid treatment

Ten slotte gaan we in deze rapportage in op een vraag die in derde meting is gesteld aan de klantmanagers over de treatments. In deze meting is gevraagd in hoeverre de klantmanagers de treatment geschikt vonden voor deze klant, met een score 0 voor geheel ongeschikt en 10 voor totaal geschikt. Hierbij moet wel weer worden bedacht dat deze vraag maar voor ongeveer de helft (N=250) van het aantal initiële deelnemers is ingevuld en we het hier over een mogelijk selectie groep hebben, namelijk de groep die niet is uitgestroomd.

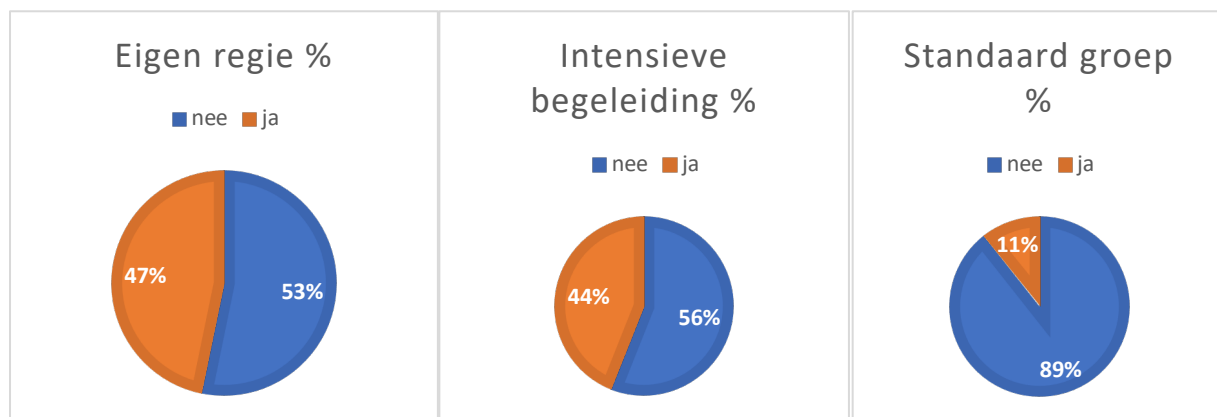
Tabel 3.2.12: Rapportcijfer voor geschiktheid treatment

Treatment geschikt (0-10)?	Gemiddelde score	N
Eigen regie	5,9	61
Intensieve begeleiding	6,3	97
Standaardgroep	7,6	92
Totaal	6,6	250

Noot: *** $p < 0,01$; ** $p < 0,05$; * $p < 0,10$, getest op basis van een multiple regressieanalyse (N=250)

Bron: Vragenlijst tweede vervolgmeting procesevaluatie Vertrouwensexperiment Apeldoorn, Tilburg University, 2020.

Tabel 3.2.12 laat zien dat de klantmanagers in de standaardtreatment het hoogste rapportcijfer geven in termen van geschiktheid, gevolgd door de intensieve begeleiding en ten slotte de eigen regie groep met een magere voldoende. Dit bevestigt het eerdere beeld dat het experiment niet per se ‘con amore’ is uitgevoerd door de consultants conform het beoogde design hetgeen deels het ontbreken van significante verschillen met de standaardgroep zou kunnen verklaren. Vervolgens is de vraag gesteld of de klantmanagers vinden dat een andere treatment meer passend was geweest voor die deelnemer. Hieronder staan de resultaten samengevat in figuur 3.2.11. Voor de klantmanagers die de eigen regie groep begeleidde, vond men voor 53% van de deelnemers deze treatment het meest passend. In bijna de helft van de gevallen had de klantmanager echter gekozen voor een andere aanpak.

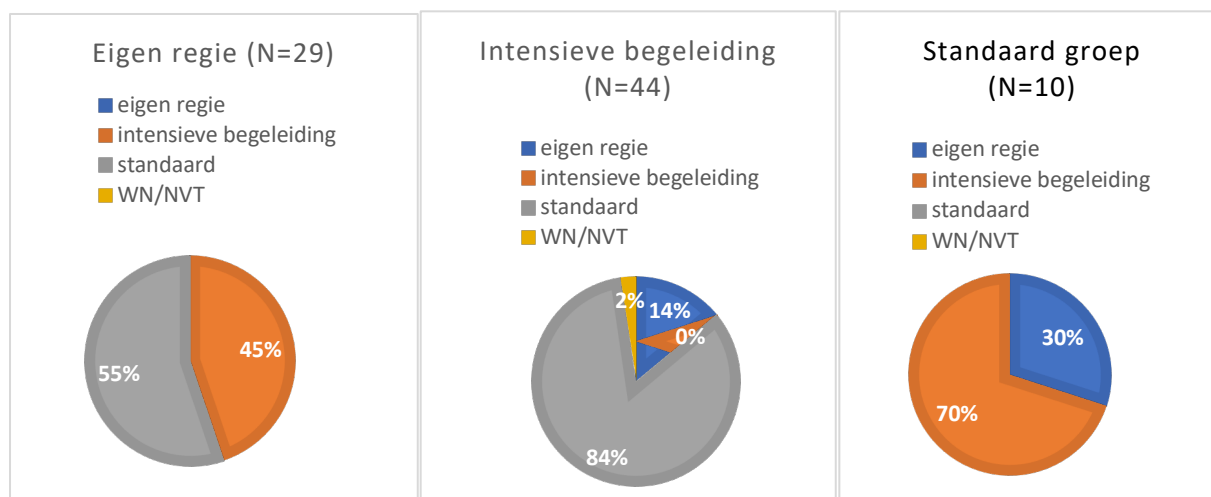


Figuur 3.2.11. Percentage ja/nee voor treatment eigen regie, intensieve begeleiding en de standaardgroep

Bij de klantmanagers van de intensieve begeleidingsgroep zien we een soortgelijk beeld ontstaan. Ook daar zouden ze in bijna de helft van de gevallen hebben gekozen voor een andere treatment. Ten slotte kijken we naar de standaardgroep en daar valt op dat slechts in 11% van de gevallen de klantmanagers een andere treatment meer passend zou hebben gevonden. Dit sluit aan bij de hoge mate van tevredenheid met de geschiktheid van deze treatment zoals uit tabel 3.2.12 bleek. De eerdere veronderstellingen over een duidelijke voorkeur bij de consultants voor de ‘klassieke’ standaard aanpak waarmee men bekend is, worden hiermee opnieuw bevestigd.

Ter verdieping van de vorige vraag is de klantmanagers ook gevraagd welke treatment ze dan zouden hebben gekozen als ze ja hadden geantwoord op de vorige vraag. Hieronder staan de resultaten samengevat.

In de 29 gevallen dat de klantmanagers van de eigen regie groep een andere treatment meer passend vindt, zou 55% kiezen voor de standaard treatment en 45% voor de intensieve begeleiding.



Figuur 3.2.12. *Indien andere treatment meer geschikt, welke is dan meer geschikt?*

Bij de 44 gevallen waarbij de klantmanager van intensieve begeleiding kiest voor een andere treatment, is het beeld minder eenduidig dan bij de eigen regie groep. Toch zou het overgrote deel (84%) kiezen voor de standaard treatment, 14% voor eigen regie en 2% weet het niet goed. Voor 10 van de deelnemers in de standaardgroep zouden de klantmanagers ook voor een andere treatment hebben gekozen. In 30% van de gevallen zou er gekozen zijn voor eigen regie en in 70% van de gevallen voor intensieve begeleiding.

Samenvattend blijkt dat de standaardtreatment het meest geschikt wordt geacht door de consultants voor een groot deel van de deelnemers. Zo blijkt dat er slechts in een klein aantal gevallen die de standaard treatment heeft gehad, er gekozen zou zijn voor een andere treatment. In de gevallen dat er begeleid is in de eigen regie of standaardgroep, dan zou er vaker gekozen worden voor de standaard treatment. Het bevestigt opnieuw het eerdere beeld. Kennelijk is de traditie in Apeldoorn/Epe er een van een stevige aanpak die men ook positief waardeert.

3.2.4 Conclusies

In deze paragraaf volgen een aantal conclusies uit de procesevaluatie. In de voorgaande paragrafen hebben we een aantal belangrijke bevindingen gerapporteerd. Ten eerste blijkt de klantmanager een belangrijk effect te hebben op de uitkomstmaten. Zo wordt in veel gevallen een positief effect van de treatments eigen regie of intensieve begeleiding ten opzichte van de standaardgroep, verklaard door het effect van klantmanager. Hierdoor vervallen de effecten van de treatments of worden ze negatief. Kort gezegd wil dit zeggen dat sommige klantmanagers van een bepaalde treatment systematisch hogere of juist lagere scores invullen dan in andere treatments, waardoor er een verschil ontstaat tussen treatments die niet toe te schrijven is aan de treatment zelf.

Vervolgens hebben we gezien dat het vaak negatieve treatment effect van eigen regie en intensieve begeleiding (in de kolom gemiddelde totaal) wegvalt op het moment dat wordt gecontroleerd voor persoonlijke kenmerken van de deelnemers. Maar welke factoren zijn dan van invloed op de uitkomstmaten? Uit onze analyses blijkt dat vrouwen en hoger opgeleiden over het algemeen hoger scoren en mensen die langer in de bijstand zitten of ouder zijn, lager scoren. Als er een onevenwichtige verdeling van deze groepen is in de treatmentgroepen, dan kan dat dus een verklaring zijn voor gevonden verschillen tussen treatmentgroepen en het blijkt dat dat hier inderdaad het geval is. Samengevat kunnen we zeggen dat de treatmenteffecten die we vonden in de kolom gemiddelde totaal, waarschijnlijk toe te schrijven zijn aan compositieverschillen tussen de groepen en dus selectiviteit binnen de groepen (ontstaan bij de aanmelding). Nader onderzoek zal nodig zijn om de precieze mechanismen te achterhalen.

We hebben ook gekeken naar de ontwikkeling van de scores op de uitkomstvariabelen over de tijd. Daar is duidelijk gebleken dat de significante positieve effecten die we in de meeste gevallen vonden, robuust en substantieel zijn. Dat betekent dat er over de groepen heen daadwerkelijk een stijging plaatsvindt over de tijd. Alleen bij het verwachte effect van de treatment en het leereffect van de deelnemer was er sprake van een negatieve ontwikkeling over de tijd. Opvallend is wel dat de standaardgroep in bijna alle gevallen een significante toename laat zien in de scores op de uitkomstvariabelen over de tijd. In een enkel geval is het positieve treatment effect niet significant, maar in de meeste gevallen wel. De standaardgroep doet het dus in de ogen van de klantmanagers erg goed. Dit is een heel verrassend resultaat, aangezien we zouden verwachten dat de standaardgroep geen effecten zou laten zien, omdat dat onze controlegroep is waarvoor niets verandert gedurende het experiment.

Deze resultaten en anderen eerder gerapporteerd doen vermoeden dat er ook binnen de standaardgroep van aanpak is veranderd. De klantmanagers van de standaardgroep draaiden ook mee in het experiment en hebben ook de bijeenkomsten bijgewoond. Daarnaast waren 2 klantmanagers actief in zowel de standaardtreatment als de intensieve begeleidingsgroep. Deze factoren hebben er waarschijnlijk voor gezorgd dat de standaardtreatment ook is veranderd gedurende het experiment, waardoor een zuivere effectmeting wordt bemoeilijkt.

Iets anders dat is opgevallen is het vaak negatieve effect van de treatments eigen regie en intensieve begeleiding over de tijd (in vergelijking met het effect van deze groepen in de nulmeting ten opzichte van de standaardgroep). Deze negatieve 'interactie-effecten' zijn robuust en blijven overeind ook met de

uitgebreide set aan controlevariabelen. Hoewel er dus geen algemene verschillen worden gevonden tussen de treatmentgroepen is er wel sprake van een verschillende ontwikkeling in de scores over de tijd voor de treatmentgroepen. Waar de standaardgroep vaak stijgt, werken de andere twee treatments vaak juist negatief uit gedurende het experiment.

Wellicht is dit te verklaren door de complexere groep deelnemers die overblijft naarmate het experiment vordert of worden de klantmanagers in hun scores beïnvloed door de mate van uitstroom die zij kunnen bewerkstelligen in de verschillende treatments. Nader onderzoek zal nodig zijn om dit helder te krijgen. Uit de focusgroep gesprekken is gebleken dat de consultants niet goed raad wisten en het instrumentarium en methodieken misten om op een goede manier met de overblijvende kwetsbare problematische groep van mensen om te gaan met het oog op begeleiding en ondersteuning waaronder veel mensen met psychische gezondheidsklachten en weinig kansen op betaald werk.

Ten slotte komt de positievere beoordeling van de standaardtreatment door de klantmanagers overeen met de toenemende scores op de uitkomstvariabelen in deze treatment. In de ogen van de klantmanagers doen deze deelnemers het goed en daardoor waarderen zij deze treatment ook goed. Het lijkt erop dat de treatments eigen regie en intensieve begeleiding op minder steun kunnen rekenen van de klantmanagers. De uitkomsten zijn wellicht anders dan verwacht, waardoor zowel klantmanager als deelnemer ontmoedigd lijkt te zijn. Aan de andere kant zouden de resultaten erop kunnen wijzen, dat indien de standaardaanpak inderdaad is veranderd, bijvoorbeeld door een andere bejegening van de cliënt, dit een effectieve methode kan zijn, die positieve resultaten oplevert voor de klantmanager en cliënt.

3.3 Focusgroep gesprekken met klantmanagers

Het focusgroep gesprek is onderdeel van de procesevaluatie welke gericht is op het proces van experimenteren met een andere benadering en uitvoering van de Participatiewet. Hoe gaan de direct uitvoerenden (klantmanagers en projectleider) maar ook de gemeente als organisatie om met de implementatie en ontwikkeling van het experiment? Het proces van experimenteren is een proces van verandering vooral ook in de benadering van de cliënt door de klantmanagers.

In deze focusgroep gesprekken wordt besproken wat de kansen en eventuele belemmeringen zijn van de andere bejegening in de relatie met de deelnemer en welke resultaten men verwacht en al dan niet bereikt. In het bijzonder wordt ingegaan op de handelwijzen van de klantmanager in de verschillende interventies en in welk opzicht ze van elkaar dienen te verschillen. Betekent 'eigen regie' dat men de deelnemer volledig moet loslaten of juist niet omdat het 'leren van zelfredzaamheid' meer contacten met de cliënt vraagt?

Ten tweede wordt ingegaan op de casuïstiek van de begeleiding waarbij specifieke problemen zoals fraude, schulden en ongemotiveerd of onaangepast gedrag aan de orde wordt gesteld om ervan te leren. Mede daarvoor is ook een training aangeboden waarbij klantmanagers onder begeleiding van een professionele coach leren hoe ze een vertrouwensrelatie met de deelnemer kunnen opbouwen en hoe ze de zelfredzaamheid van deelnemers kunnen bevorderen.

Tot slot wordt ingegaan op de context waarbinnen het vertrouwensexperiment plaatsvindt, dat is de gemeentelijke organisatie, de politiek, de media en alle betrokkenen daaromheen. Deze gesprekken

vonden plaats na 6 tot 9 maanden in het bijzijn van de projectleider en de strategische beleidsadviseur die het experiment beleidsmatig en uitvoeringstechnisch mede coördineert.

In totaal zijn er 4 focusgroep gesprekken gehouden. Het eerste gesprek vond plaats op 18 januari 2018 nadat het experiment enkele maanden van start was gegaan en het eerste gesprek met de cliënt is gehouden. Het tweede gesprek heeft op 27 september 2018 plaatsgevonden dus na bijna 1 jaar en het derde gesprek volgde op 20 mei 2019. Het vierde en laatste focusgroepgesprek is gevoerd op 6 november 2019.

3.3.1 Onderwerpen

Tijdens de focusgroep gesprekken is gebruik gemaakt van een PowerPointpresentatie als leidraad voor het gesprek. In deze PowerPointpresentatie werden 5 stappen onderscheiden die ontleend zijn aan de stappen in het proces van experimenteren. Deze 5 stappen zijn:

- Stap 1. De beleidsuitgangspunten en het doel van het experiment
- Stap 2. Achterliggende theorie/filosofie en wetenschappelijke vragen
- Stap 3. Van theorie naar RCT-design voor een niet-AMvB experiment
- Stap 4. Van design naar implementatie: de drie treatment groepen; de relatie met de cliënt; de eigen rol als coach etc.
- Stap 5. Van implementatie naar uitkomsten: de lokale en politieke context, de randvoorwaarden en de uitkomsten voor de cliënt.

Aan de hand van deze stappen zijn er in de PowerPointpresentatie topics opgenomen die meer inzicht bieden in de ervaring van de klantmanagers met deze stappen.

3.3.2 Analysemethode

Elk van de vier focusgroep gesprekken zijn door middel van een mobiele telefoon opgenomen, met toestemming van de deelnemers aan het gesprek. Na opname zijn deze gesprekken anoniem getranscribeerd en vervolgens bewerkt tot een samenhangend verslag waarin de structuur van de stappen is aangehouden. In de praktijk lopen de stappen echter tijdens het gesprek soms door elkaar heen, waardoor bepaalde onderwerpen op meerdere plaatsen terug kunnen komen. Vervolgens zijn de verslagen van deze drie focusgroep gesprekken met elkaar vergeleken, met als uitgangspunt de structuur van de verschillende stappen en is er een analyse gemaakt van wat de standpunten op deze stappen waren en hoe deze veranderden gedurende het proces. Ook is er in de analyse nadrukkelijk aandacht besteed aan zogeheten 'implementatie fouten', wat inzicht geeft in de mate waarin de treatments daadwerkelijk zijn uitgevoerd, zoals ze theoretisch bedoeld waren ('theory failure'), of dat de praktijk van uitvoering er anders uit zag, wat dan de implementatie fouten zijn ('implementation failure').

3.3.3 Resultaten

Tijdens de vier focusgroep gesprekken die zijn gehouden, zijn er verschillende stappen besproken. In praktijk valt wat wordt besproken vaak onder meer dan een stap, omdat de onderwerpen elkaar deels overlappen. In deze analyse is er voor gekozen toch zoveel mogelijk de opgenomen gesprekken te beschrijven

Stap 1. De beleidsuitgangspunten en het doel van het experiment

In deze eerste stap is besproken of de klantmanagers de uitgangspunten van het experiment en het doel helder voor ogen hebben. Daarnaast is er gevraagd of in de praktijk duidelijk wordt welke aanpak, belonen of straffen, beter werkt en voor welke groepen. Tijdens het eerste focusgroepgesprek met de klantmanagers, werd al duidelijk dat het experiment brede steun krijgt wat betreft de uitgangspunten, de filosofie en de uitwerking. Voor een deel is dit te verklaren doordat klantmanagers zichzelf actief hebben aangemeld voor deelname aan het experiment en doordat een klein aantal klantmanagers, die zich er minder in konden vinden, het experiment hebben verlaten.

De klantmanagers geven aan tijdens het tweede focusgroepgesprek dat het grote verschil met de reguliere groep is dat vertrouwen echt centraal staat bij de experimentele groepen. Dit benadrukt het feit dat het uitgangspunt voor de experimentele groepen wezenlijk anders is dan voor de standaardgroep en dat dat ook zo wordt ervaren door de klantmanagers. Tijdens het derde focusgroepgesprek blijkt dat de werking van de treatments om de gestelde doelen te bereiken, sterk afhangt van de persoon van de deelnemer in begeleiding. Binnen de eigen regie groep blijken nieuwe instromers vaak beter in staat te zijn zelf een plan te maken en regie te pakken, terwijl de cliënten die na 1,5 jaar nog in het experiment zitten daar niet toe in staat is en ook moeilijk te motiveren en te contacteren is.

Stap 2. De achterliggende theorie

In deze tweede stap lag de focus op de achterliggende theorie van het experiment en hoe dat werkt in de praktijk. Op deze stap is met name in het eerste en tweede focusgroep gesprek ingegaan. Tijdens het eerste focusgroepgesprek blijkt dat klantmanagers vooral enthousiast zijn over de focus op een andere bejegening van de cliënt en de oriëntatie op de intrinsieke motivatie van de cliënt. Volgens hen zijn deze twee theoretische uitgangspunten bepalend voor het succesvol kunnen bemiddelen en begeleiden naar werk of naar andere participatiemogelijkheden. Zo wordt opgemerkt dat de cliënt “zijn of haar eigen ideeën en dromen beter kan vormgeven”. Ondanks deze positieve inzichten, wordt in de dagelijkse praktijk duidelijk dat het opbouwen van een vertrouwensrelatie met de cliënt lastig is en tijd vraagt door het grote wantrouwen van de cliënt naar de gemeente toe. Zo zijn cliënten in eerste instantie achterdochtig en wantrouwend en vragen zich openlijk af of de andere bejegening wel ‘echt’ is. Later tonen ze zich verbaast over de andere houding en gedrag van de klantmanager. Veel cliënten tonen ook hun emoties en huilen zelfs van dankbaarheid voor de hulp en begeleiding en vooral de andere bejegening die hen ten deel valt.

Ook tijdens het tweede focusgroepgesprek geven de klantmanagers aan dat cliënten positief reageren op de begeleiding op basis van vertrouwen en dat ze de extra tijd voor elke cliënt en dus voor het ‘echte’ gesprek alsook de relaxtheid in de begeleiding (ruimte voor de cliënt) positief waarderen. Door de klantmanagers wordt dit dan ook voor de toekomst als een aantrekkelijk en houdbaar model gezien, maar men is zich meer rekenschap gaan geven van de lastige vraagstukken en dilemma’s die daarbij om de hoek komen kijken. Echter, de gemeentelijk organisatie is nog steeds erg gericht op het resultaat van voltijdse uitstroom naar betaald werk, wat niet overeenstemt met de uitgangspunten en de onderliggende theorie van het experiment, waarin het proces van begeleiden belangrijker wordt geacht voor duurzame uitstroom dan het korte-termijn resultaat in de vorm van voltijdse uitstroom naar werk.

Stap 3. Het design van de treatments

Tijdens de gesprekken is er in deze stap aandacht besteed aan het design van de verschillende treatments en de mate waarin deze treatments op de beoogde manier in de praktijk uitvoerbaar zijn.

Onderscheidend vermogen van de treatments

In Apeldoorn en Epe is er in eerste instantie voor gekozen om klantmanagers verantwoordelijk te maken voor een enkele treatment en niet voor meerdere treatments, maar uiteindelijk hebben toch twee klantmanagers zowel de standaardgroep begeleid als de intensieve begeleidingsgroep. De treatments zijn van tevoren zo goed mogelijk op papier gezet zodat iedereen het zoveel mogelijk op dezelfde manier doet. Hierdoor is de kans op vermenging en het afnemen van het onderscheid tussen de treatments kleiner.

Desondanks blijkt het in de praktijk vaak lastig voor klantmanagers om te bepalen of hun gedrag binnen de uitgangspunten en de filosofie van het experiment past. Ze zijn daarover onzeker en er komt dan ook naar voren dat zij behoefte hebben aan een handelingskader. Apeldoorn en Epe hebben daartoe naast de zelfredzaamheid training ook intervisie ingebouwd om van elkaar te leren en dit kader gemeenschappelijk te gaan ontwikkelen. Men ervaart met andere woorden een gebrek aan handvatten of methodieken om de deelnemer op deze andere manier goed te kunnen ondersteunen.

De gehanteerde methodiek die onderwerp was van de zelfredzaamheid training (gegeven door Prof. Roland Blonk) wordt als waardevol ervaren door de klantmanagers en van belang om de interventies op een goede manier te begeleiden. Echter, er was wel twijfel over of deze methodiek voldoende differentiateerde tussen de treatments. Ook geven klantmanagers aan te leren gedurende het experiment, waardoor de eigen aanpak verandert om deze effectiever te maken. Om die reden is vergelijking met een referentiegroep die niet deelneemt aan het experiment en het onderzoek, en dus minder verandert, van belang. Dat is evenwel voor de uitkomstmaten in de procesevaluatie (zoals zelfredzaamheid bij deelnemer) niet mogelijk maar wel voor de uitkomstmaat voltijdse uitstroom naar betaald werk in de productevaluatie (door vergelijking met de gegevens in de bijstandsuitkeringsstatistiek). De klantmanagers geven aan dat de begeleiding verschilt en dus afhankelijk is van de klantmanager, ook bij de standaardtreatment. Dat laatste heeft ook te maken met beleidsveranderingen en vooral een ingrijpende en langdurige reorganisatie van de dienstverlening gedurende de looptijd van het experiment (overgang van klantmanagers naar trajectregisseurs). Ondanks de door de klantmanagers ondersteunde op vertrouwen gebaseerde aanpak van de treatments laten zij in Apeldoorn/Epe zien dat bij moeilijke cases er soms wordt teruggevallen op de klassieke 'stok en wortel' benadering en daarbij vooral de 'stok', in de vorm van het opleggen van maatregelen en het strakker sturen op trajecten en activiteiten, bijvoorbeeld omdat een deelnemer niet op oproepen reageert. In die gevallen wordt er afgeweken van de uitgangspunten van het experiment, waarin wordt gesteld dat de klantmanager ten alle tijden in gesprek moet blijven en contact moet proberen te zoeken met de deelnemer en zodoende er alles aan moet doen om niet tot een maatregel over te gaan omdat wanneer dat gebeurt de moeizaam opgebouwde vertrouwensrelatie wordt verbroken. Kennelijk ervaren klantmanagers in dergelijke moeilijke gevallen onvoldoende handvatten of methodieken om vanuit vertrouwen te kunnen blijven opereren en vallen ze terug op het 'bekende' reguliere recept.

Voor- en nadelen per treatment

Over het algemeen geven de klantmanagers aan dat ze niet denken dat de ene treatment beter werkt dan de andere en dat het juist maatwerk blijft. Volgens de klantmanagers is er wel een breder en beter gesprek mogelijk waarbij niet alleen werk maar ook andere zaken die het leven van de mensen raken aan de orde kunnen komen.

- **Eigen regie**

Uit de gesprekken blijkt dat klantmanagers die betrokken zijn bij de 'eigen regie' treatment soms problemen hebben om de grens te bepalen tussen 'leren zelfredzaam te worden', hetgeen aan het begin meer begeleidingstijd vraagt en 'afstand nemen' of zo veel mogelijk loslaten en het aan de cliënt overlaten om actie te nemen ('zelf aan het stuur'). Het 'loslaten' en afhouden van contact zou het vertrouwen met de cliënt kunnen schaden, omdat de cliënt vraagt om en behoefte heeft aan contact en hulp.

Het "loslaten" vraagt ook veel van de klantmanager, omdat zij gewend waren om de 'regie' te pakken en mensen op weg te helpen in plaats van de regie bij de mensen zelf te leggen (zelfredzaamheid). Vaak is het ook moeilijk om contact te krijgen met de deelnemer zeker wanneer de deelnemer zelf geen contact opneemt. Tijdens het tweede gesprek wordt de verwachting geuit dat dit probleem alleen maar groter zal worden omdat de groep die overblijft lastig in beweging te krijgen is en ook niet zelf met een plan kan komen. Sommige cliënten daarentegen interpreteren eigen regie als niets meer hoeven doen en achterover leunen, waardoor het idee van 'zelfredzaamheid leren' niet meer mogelijk is. Ook zijn er veel mensen met psychische problemen of een ernstige schuldproblematiek, die eigenlijk meer baat hebben bij een intensieve begeleiding in plaats van eigen regie, welke voor hen niet de meest passende aanpak is. Onder stap vier wordt verder besproken hoe het uitvoeringsproces binnen de treatment eigen regie is verlopen. De klantmanagers geven wel aan dat het bij eigen regie fijn is dat er niet zoveel gecontroleerd hoeft te worden.

- **Intensieve begeleiding**

Bij intensieve begeleiding zijn er wisselende resultaten: er zijn zowel moeilijke gevallen, als cliënten die vrij snel stappen maken en aan het werk gaan. De intensieve treatment blijkt ook intensief te zijn voor de klantmanager: er moeten veel gesprekken worden gevoerd die soms als niet zinvol worden ervaren. Soms ervaren klantmanagers hun eigen rol dan meer als maatschappelijk werker of psycholoog welke rol zij ervaren als een andere rol dan die van hen wordt verwacht. Zij hebben daarvoor ook niet de kennis en vaardigheden waardoor verwijzing naar andere hulpverleningsinstanties nodig is. Aan de andere kant geven deze gesprekken wel de mogelijkheid om te werken aan kleinere (sub)doelen waar ook stappen in gemaakt kunnen worden. Het intensieve contact met de deelnemer wordt als prettig ervaren en ook goed voor hen omdat er ook af en toe minder serieuze zaken besproken kunnen worden en daardoor er een vertrouwensband kan worden opgebouwd.

In het derde focusgroepgesprek komt naar voren dat de groep die is overgebleven bij deze treatment, een groep is met weinig zelfdiscipline om de eigen ideeën uit te voeren en die slecht met tegenslagen om kan

gaan. Vaak zijn dit mensen met een psychische of sociale problematiek met een grote afstand tot de arbeidsmarkt. Mensen lijken soms ook niet te begrijpen waarom ze zo vaak worden opgeroepen (6 keer in principe per jaar) en dan is het ook voor de klantmanager lastig om de positieve 'flow' erin te houden en een zinvolle ondersteuning te kunnen bieden. De eerste groep die is uitgestroomd bleek wel in staat te zijn zelf actie te ondernemen. Volgens een klantmanager wordt na een maand of 3 in de begeleiding duidelijk of iemand het zelfstandig kan en initiatief toont of niet.

- **Standaard begeleiding**

De begeleiding van de standaardgroep is, vergeleken met de andere treatments, relatief eenvoudig aangezien de manier van begeleiding veel meer vastligt en in Apeldoorn/Epe redelijk strikt wordt toegepast. Het onderscheid tussen kansrijke cliënten die bij Lucrato worden begeleid en kansarmere cliënten die bij Activering en Inkomen worden begeleid helpt daarbij. Zeker bij Lucrato wordt er een strikt kader gehanteerd om cliënten te begeleiden. De druk vanuit de organisatie is onverminderd groot om voltijdse uitstroom naar betaald werk te realiseren.

Voor welke groepen werkt de aanpak minder goed?

Tijdens de gesprekken geven sommige klantmanagers aan dat het sterk afhangt van de persoon of een begeleiding aanslaat of niet. Veel deelnemers hebben gezondheidsproblemen, waardoor ze niet kunnen werken, maar soms is dat volgens sommige klantmanagers ook een middel om mee te schermen om niet te hoeven werken. In dat geval komt er een Arboarts aan te pas. Vooral bij de eigen regie groep is het dan lastig als mensen wel blijken te kunnen werken, omdat ze niet gedwongen kunnen worden. Ook hebben mensen die vaak lang in de bijstand zitten geen perspectief richting betaald werk. Hier moet echter rekening gehouden worden met dat het niet alleen gaat om uitstroom naar betaald werk, maar ook om een bezigheid, vrijwilligerswerk of therapie.

Van de deelnemers die in de laatste fase van het experiment nog in begeleiding zijn, zijn er veel mensen met een psychische of sociale problematiek, zoals laagopgeleide oudere mensen met veel schulden en/of ernstige medische problemen (ziekten). Soms zijn dit volgens sommige klantmanagers ook mensen die graag in de bijstand willen blijven, vanwege hun medische problemen (werk is geen optie) en de veiligheid die het hen biedt. In stap 4 wordt dit uitgebreider besproken.

Stap 4. De treatment groepen, de relatie met de cliënt en de rol van de klantmanager

In stap 4 is een groot aantal onderwerpen besproken. Er is besproken of de treatment in de praktijk effectief is voor de deelnemers, hoe klantmanagers hun eigen rol zien, welke positieve en negatieve ervaringen de klantmanager heeft met het uitvoeren van de treatments, welke behoefte er is aan ondersteuning, de verwachtingen ten aanzien van het management en de organisatie, de eigen motivatie en het werkgeluk van de klantmanager zelf alsook de leereffecten bij deelnemer en klantmanager. In alle focusgroep gesprekken is hier aandacht aan besteed, maar voornamelijk in het derde en vierde gesprek.

Effectiviteit treatment

Tijdens het derde gesprek is er met de klantmanagers gesproken over hoe zij denken over het effect van de treatment op de deelnemer. Uit de procesevaluatievragenlijsten werd duidelijk dat de klantmanagers

niet erg positief zijn over de ontwikkeling van cliënten met betrekking tot zelfredzaamheid en zelf in actie komen. Ook de mate van realisme van de eisen die cliënten stellen aan een toekomstige baan leek achteruit te gaan. De klantmanagers geven aan dat dat inderdaad zo is. De lage mate van realisme zou volgens hen ook te maken hebben met de doelgroep waarmee ze werken, omdat die weinig reflecteert op zichzelf. Zij willen bijvoorbeeld een droombaan die onbereikbaar is en sommige cliënten hebben meer hulp nodig dan de klantmanagers kunnen bieden. Daarentegen lijken veel klantmanagers van oordeel dat hun cliënten stijgen op de participatieladder. De klantmanagers beamen dit. Sommige cliënten met een grote afstand tot de arbeidsmarkt komen toch in beweging en gaan activiteiten ondernemen zoals vrijwilligerswerk en dagbesteding. De klantmanagers zien dit als een gevolg van de aandacht die de mensen in de treatments krijgen.

Ondanks de verbeterde economische situatie, die meer perspectief biedt voor de deelnemers op de arbeidsmarkt, blijkt er toch vaak een drempel voor mensen te bestaan om aan het werk te gaan. De bijstand is een vorm van veiligheid en zekerheid en ze weten volgens de klantmanagers dat als ze uitstromen, ze ook niet zo gemakkelijk weer kunnen instromen in de bijstand. Daarnaast zien mensen op tegen alle onzekerheden die erbij komen kijken als ze aan het werk gaan (bv. toeslagen).

In het vierde focusgroepgesprek wordt duidelijk dat een aantal deelnemers uit de intensieve begeleidingsgroep in het experiment in het laatste halfjaar alsnog in beweging is gekomen. Door de klantmanagers werd geopperd dat zodra voor hen duidelijk werd dat het experiment op korte termijn zou gaan stoppen en zij daarna weer onder de reguliere dienstverlening zouden komen te vallen, ze zelf weer het initiatief namen om werk te vinden. Door anderen werd geopperd in het focusgroepgesprek dat door het ‘achterover leunen’ in de laatste fase bij gebrek aan mogelijkheden voor begeleiding de deelnemers alsnog in beweging kwamen maar wel geholpen door de gunstige arbeidsmarkt. Per 1 oktober 2019 is het experiment zoals beoogd beëindigd. De indruk bij sommige klantmanagers was dan ook dat deze deelnemers wellicht eerder werk hadden kunnen vinden, omdat ze fysiek verder in orde waren. Volgens hen is het voor sommige cliënten ook prettig om toch een stok achter de deur te hebben. De meeste cliënten die niet in beweging komen geven aan dat ze ziek zijn. Volgens sommige klantmanagers is dat controleerbaar, maar soms ook niet en dan lijkt het alsof ze ermee schermen om niet aan het werk te gaan. In het gesprek werden door de klantmanagers een aantal barrières genoemd die ervoor kunnen zorgen dat cliënten niet in beweging komen:

- Mensen zitten langdurig in de bijstand en hadden geen contactpersoon en weinig gesprekken bij de gemeente. Ze zijn vergeten volgens de klantmanagers.
- In Epe zou het er ook aan liggen dat dat een dorp is. Mensen moeten dan met de fiets of met de bus naar het werk en dat houdt hen tegen. Als ze met de bus moeten valt heel veel productiewerk weg, want dat is allemaal ploegenwerk. Van Apeldoorn naar Epe is een halfuur/ drie kwartier fietsen en het werk is vaker in Apeldoorn dan in Epe, waardoor mensen uit Epe de afstand als een drempel kunnen ervaren.
- Sommige mensen zijn ziek, sommigen zijn verslaafd of eenzaam.
- Voor sommige cliënten die vrijwilligerswerk doen, is dit vrijwilligerswerk vergelijkbaar met een betaalde baan, qua tijd en verplichtingen die het met zich mee brengt. Daar voelt men zich veilig en gewaardeerd en wil het daarom niet loslaten. Betaald werk kunnen ze zo weer kwijt zijn en

daarom durven ze het risico niet te nemen, want als dat gebeurt, dan hebben ze geen baan meer en geen vrijwilligerswerk meer.

- Sommige mensen hebben schulden die afbetaald moeten worden. Voor hen loont het weinig om te gaan werken, omdat het loon dan eerst naar de schuldeisers gaat.
- Sommige cliënten hebben nooit gewerkt, bijvoorbeeld vrouwen met kinderen die nu volwassen worden. Die mensen zijn altijd met rust gelaten door de gemeente, omdat ze moesten zorgen maar daardoor is er niet gewerkt aan economische zelfstandigheid. Het is dan heel moeilijk om na al die tijd werk te vinden.
- Voor alleenstaande moeders loont het niet om (meer) te gaan werken. Als ze al parttime aan het werk zijn is het nog moeilijker om ze in beweging te krijgen, want tussen de 12 en 20 uur werk loont het nog om in de bijstand te blijven. Daarnaast krijgen ze te maken met afbouw van toeslagen, die minder worden of verdwijnen als ze (meer) gaan werken.
- Sommige cliënten zouden al veel eerder in het Doelgroepen Register opgenomen moeten zijn, omdat ze een bepaalde problematiek hebben die eerder over het hoofd is gezien en daardoor een drempel vormde om aan het werk te gaan. Door opname in het register, kunnen ze activiteiten gaan zoeken die wel passend zijn, bijvoorbeeld parttime werk. Echter, het aanvragen van een registratie is een vrije keus en kan niet worden opgelegd aan een cliënt. Sommige cliënten willen daar niet aan meewerken.

Over de alleenstaande moeders wordt nog uitvoeriger doorgepraat. Aan de ene kant kan het ook als waardevol worden beschouwd dat alleenstaande moeders die parttime werken een gezin hebben dat in balans is, waarvan de kinderen naar school gaan en er geen gedoe is in huis. Daarnaast verdienen ze een deel van de uitkering terug. Aan de andere kant is het wel een beleidskwestie, want dan moet men accepteren dat iemand bijvoorbeeld maar 15 uur per week werkt. Zou er van een alleenstaande moeder ook niet verwacht kunnen worden dat ze bijstandsniveau verdient, dus zo'n 25 uur per week werkt? In dat geval zou er in de landelijke politiek op beleidsterrein wel wat aan de afbouw van toeslagen moeten worden gedaan. Sommige deelnemers stromen pas uit de bijstand als hun kinderen boven de 18 zijn, omdat dat gevolgen heeft voor de toeslagen. Een bepaalde doelgroep blijkt dan slim aan het rekenen te zijn geweest volgens de klantmanagers, die zouden wel meer uren kunnen werken, maar doen het om financiële redenen niet.

Ook over het laatste punt, het Doelgroepen Register, wordt nog verder doorgepraat. Er wordt aangegeven dat het ook vanuit het gemeentelijk bestand niet duidelijk is dat iemand een registratie in het Doelgroepen Register heeft. Dat zorgt ervoor dat de beperking van een cliënt niet voldoende op het netvlies is tijdens de intake of tijdens de gesprekken, terwijl het voor die mensen wel iets kan opleveren. Pas als er voldoende tijd en aandacht is voor de cliënt, dan kunnen deze zaken beter worden uitgevraagd, volgens de klantmanagers. Dat dit te weinig is gebeurd heeft volgens de klantmanagers ook te maken met de 'targets' die gehaald moeten worden, waardoor ze te weinig tijd hebben om iemand te leren kennen. Op het moment dat de druk erop staat om doelen te halen en iemand komt te boek te staan als 'niet-willer' dan wordt hij/zij snel naar achter geschoven en vergeten of iemand wordt een kant op geduwd die misschien niet passend is bij de beperking van die persoon. De klantmanagers kennen verschillende cases

waar iemand als 'niet-willer' te boek staat, maar proberen nu wel meer om erachter te komen waarom iemand niet wil, bijvoorbeeld door zo iemand naar het diagnosecentrum te sturen.

Meer dan dertig procent van de mensen in de bijstand hebben psychische klachten en combineren dat met fysieke gezondheidsklachten. Een meerderheid geeft aan gezondheidsproblemen te ervaren. Soms lijken deelnemers daarmee te schermen. Toch heeft het gemeentelijk beleid daar ook aan bijgedragen met de zorg- en hulpverleningstrajecten, waardoor bijstandsgerechtigden in een 'ik kan niet werken'-stand werden gezet. Nu is er meer kennis over dat zorg, werk en herstel ook samen kan gaan. Daar valt wel nog veel in te verbeteren. Een analyse van het bestand in Apeldoorn wijst uit dat 70% van de deelnemers vindt dat ze psychisch niet gezond genoeg zijn om te werken en dat veel mensen zowel psychische als fysieke problemen hebben. In Apeldoorn lijkt dat erger te zijn dan in enkele andere gemeenten. Een gedachte die wordt geopperd is dat heel veel mensen die nooit gewerkt hebben, maar wel medische klachten hebben, nooit via het UWV-kanaal in de WIA terecht zijn gekomen. Op die manier blijven deze mensen dus eigenlijk als verkapte WIA-deelnemer in het bestand zitten, maar wordt de benadering daar niet op aangepast. Er worden door de gemeente 180 arbeid medische onderzoeken per jaar gedaan hetgeen een signaal zou kunnen zijn, dat er veel mensen per jaar zo'n traject in gaan.

Tijdens dit gesprek komt men tot de conclusie dat de heterogeniteit van het bestand en de hoeveelheid verschillende problematieken ervoor zorgt dat maatwerk bijna onmisbaar is. Hierdoor is het niet eenvoudig al van tevoren een idee te hebben wat de beste methode is. Juist door tijd en aandacht en meer rust in te bouwen, kan er voor het merendeel van de gevallen meer worden bereikt. Dat is de opbrengst van het experiment. Voor een ander deel van de cliënten werkt het minder goed en had er wellicht meer kunnen worden bereikt met meer sturing door de klantmanager. De werkbare elementen van het experiment zijn volgens de klantmanagers het wederzijds vertrouwen, het gesprek aangaan met de deelnemer en meer vanuit de intrinsieke motivatie stappen zetten. Maar er zou wel op een gegeven moment meer sturing mogen komen om sommige cliënten over een bepaalde drempel te helpen waar ze zelf niet overheen kunnen. Maar ook dat blijft maatwerk voor wie dat wel of niet werkt.

Behoefte aan training/intervisie

De gemeenten Apeldoorn en Epe hebben gekozen voor een training en intervisie sessies voor de klantmanagers om de methodiek eigen te maken en probleemgevallen te kunnen bespreken en een gezamenlijke aanpak te vinden. Het schriftelijk vastleggen van de inhoud van de treatments heeft ertoe bijgedragen dat men bewust wordt van de grenzen van de verschillende treatments en dat men dus weet of men in de praktijk binnen of buiten de lijntjes kleurt. Echter, elke 'case' is specifiek en vraagt een individuele aanpak hetgeen soms lastig is. Dit vraagt daarom voortdurend afstemming en aanpassing hetgeen de betrokkenen kenschetsen als een leerproces van vallen en opstaan. Ook geven de klantmanagers aan dat zij behoefte hebben aan meer intervisiemomenten om niet in de oude werkwijze te vervallen. In de praktijk is het lastig om de treatments in het experiment zuiver te houden. Daarnaast opperen de klantmanagers dat het goed zou zijn dat een klantmanager na 2 jaar een bestand moet overdragen aan een collega, zodat de deelnemer de kans krijgt op een ander gesprek.

Ook geven de klantmanagers aan dat er soms situaties ontstaan waarin duidelijk wordt dat de deelnemer eigenlijk een andere treatment nodig heeft, maar er toch vastgehouden moet worden aan de treatment

waarin iemand is ingedeeld. Dit kan zorgen voor een professioneel rolconflict, waardoor het voor de klantmanager een uitdaging is hiermee om te gaan en de treatment zuiver te houden.

Ten slotte geven de klantmanagers in het derde gesprek aan dat ze meer specifieke intervisie sessies zouden willen hebben over bijvoorbeeld autisme of andere psychische problematieken. Het is echter wel de vraag in hoeverre een psychosociale begeleiding echt past bij de functie van klantmanager. Dit punt komt in een later gesprek terug en wordt verderop meer uitvoering besproken.

Leereffecten en rol van de klantmanager

In het derde gesprek geven de klantmanagers aan dat de manier waarop ze begeleiden verschilt van hoe ze aan het begin van het experiment begeleidden. Ze geven aan dat ze veel meer spiegelen met de deelnemer over wat er moet gebeuren. Daarnaast is het wel moeilijk om mensen te blijven begeleiden die nu al het maximaal haalbare doen. Ook ervaren ze dat het fijn is om de actie bij de ander neer te kunnen leggen. Het effect is misschien hetzelfde, maar het is wel fijner werken voor de klantmanagers. Klantmanagers zijn minder gaan sturen. Wat betreft de rol van de klantmanager, geven sommige klantmanagers aan dat ze vaak gesprekken voeren als een soort maatschappelijk werker. Dat is niet erg, zolang er tijd voor is, maar soms moeten mensen ook worden doorverwezen naar andere hulpverlening. Toch is het ook afhankelijk van de klantmanager en de deelnemer welk gesprek er wordt gevoerd.

Tijdens het vierde gesprek wordt besproken hoe de begeleiding in met name de eigen regie groep is verlopen en wat daar de leereffecten zijn geweest. Daar blijkt dat er geen sprake is geweest van 'loslaten', maar juist van intensieve aandacht en veel tijd voor de deelnemer aan het begin van de treatment. De kennismakingsgesprekken zijn uitvoerig geweest en daarin is aandacht besteed aan allerlei onderwerpen. Soms is het onderwerp 'werk' niet eens besproken in dat gesprek. Het opbouwen van de relatie wordt als belangrijkste element gezien om zo later de gesprekken makkelijker te kunnen oppakken. Soms komen mensen alleen al daardoor zelf in beweging. De intensiteit van de begeleiding zat aan het begin van het experiment en werd daarna organisch minder, dan mochten cliënten zelf contact opnemen. Achteraf gezien heeft de treatment eigen regie wel wat last gehad van de manier waarop eigen regie aan de start van het experiment was ingestoken, namelijk de burger neemt contact op wanneer hij/zij wil, zoals ook eerder werd besproken onder stap 3. Dat is heel anders dan wanneer de afspraak is: de gemeente bepaalt wel en nodigt je uit en je komt gewoon, maar je hebt eigen regie over het resultaat. Op die manier hebben de klantmanagers het ook beleefd. Hierin verschilt het experiment in Apeldoorn ook van dat van bijvoorbeeld Groningen, waar eigen regie echt loslaten was. Om cliënten eigen regie te leren is volgens de klantmanagers contact nodig, anders is er volgens hen ook niet zoveel verschil met een standaardaanpak waarbij mensen jarenlang niet gesproken worden.

Een ander belangrijk leerpunt vinden de klantmanagers dat je juist de tijd en aandacht moet richten op cliënten die langer dan 5 jaar in de bijstand zitten, omdat je dan anders gaat acteren en een heel ander gesprek voert. Ook zouden klantmanagers gemotiveerd kunnen worden door doelen te stellen die haalbaar zijn, waardoor het werkgeluk ook omhooggaat. Als de lat te hoog ligt, gaan klantmanagers keuzes maken en wordt er al snel gekozen voor mensen die makkelijk bemiddelbaar zijn. De tijd en begeleiding moet aan mensen worden gegeven die het hard nodig hebben en niet aan mensen die aan de voorkant al plaatsbaar zijn, volgens de klantmanagers.

Rapportcijfer

In het vierde gesprek is de balans opgemaakt en werd de klantmanagers gevraagd om een rapportcijfer te geven aan het experiment. De klantmanagers geven unaniem een 8 voor het experiment zelf, voor de waarde ervan voor de deelnemers en voor henzelf als klantmanagers. Over de uitvoering ervan is niet iedereen even tevreden, bijvoorbeeld omdat het invullen van de vragenlijsten veel tijd kostte en het lastig was om deelnemers te bewegen de vragenlijst in te vullen, waardoor de respons niet heel hoog is. Dit ligt waarschijnlijk aan het feit dat respondenten de vragenlijsten niet meer invullen als ze zijn uitgestroomd. Bij sommige andere gemeenten was er een stok achter de deur, waardoor er een hogere respons bereikt is. Ook blijkt dat veel oudere deelnemers moeite hadden de vragenlijst in te vullen.

Het experiment heeft voor de klantmanagers opgeleverd dat het leuker is om gesprekken op deze manier te voeren en dat het meer rust oplevert voor de deelnemer en voor henzelf. Daarnaast zien ze ook een beter resultaat. Door de deelnemer mee te laten denken creëert de klantmanager ook meer draagvlak voor wat hij doet en het wordt echt gewaardeerd door de klant.

De klantmanagers zijn wel sceptisch over wat dit experiment voor uitkomst zal hebben voor de wetgeving. Zolang de politiek 'rechts' georiënteerd is, zal er minder draagvlak zijn voor een sociaal in plaats van economisch mensbeeld binnen het beleid. Van belang zal vooral zijn hoe de Tweede Kamer er tegenaan kijkt. Ook hebben de klantmanagers hun twijfels omdat de verschillende aanpakken niet zozeer veel verschil in uitstroom lijken te genereren. Uit de voorlopige resultaten in een andere gemeente lijkt er wel een substantieel en significant verschil te zijn tussen de treatments en referentiegroep, maar niet zozeer met de standaardgroep. Dat komt waarschijnlijk doordat de standaardtreatment mee veranderd is met het experiment. Verder blijkt dat de rol van de klantmanager belangrijk is geweest. In veel gevallen heeft de klantmanager niet zo heel veel effect, wat goed is, maar in enkele gevallen wel en dan kan het effect bijvoorbeeld ook sterk negatief zijn. Voor de vorming van beleid is dit heel belangrijk, omdat het kan helpen om bepaalde keuzes te maken, maar ook om te kijken hoe klantmanagers gefaciliteerd kunnen worden (methodiekontwikkeling) om hun werk goed te doen.

Stap 5 de lokale en politieke context, de randvoorwaarden en de uitkomsten voor de deelnemer

In stap 5 is er gesproken over de context waarbinnen het experiment plaatsvond. De omstandigheden waarin klantmanagers hun werk deden, bijvoorbeeld de case-load maar ook de interne organisatie van het experiment en de rol van het management. Daarnaast welke rol de politieke en lokale context heeft gespeeld en welke beleidsaanbevelingen er gedaan kunnen worden.

Caseload

Tijdens het eerste gesprek geven de klantmanagers aan de nieuwe manier van begeleiding als relaxter en prettiger te ervaren omdat er minder caseload is. Er is meer tijd voor het 'echte gesprek' met de cliënt, waardoor er een vertrouwensband ontstaat, die kan helpen om effectiever te begeleiden. Echter, de klantmanagers twijfelen of de extra begeleidingstijd zich altijd uitbetaalt in effectiviteit. Vooral de individuele situatie en problematiek van de cliënt speelt daarbij een belangrijke rol, maar ook de onzekerheid en het vertrouwen van de klantmanager zelf of de nieuwe aanpak echt gaat werken voor deze persoon.

De caseload is afhankelijk van het type treatment naar beneden bijgesteld, maar werd nog steeds als belastend ervaren door sommigen. Een caseload tussen de 75 en 90 is behapbaar en boven de 90 is (te) veel, maar mogelijk wel haalbaar in de standaardgroep. De caseload was een voortdurend aandachtspunt van bespreking tussen de coaches. Tijdens het derde gesprek blijkt dat de klantmanager met een caseload van 80 toch te weinig tijd heeft om de cliënten gedurende de week te volgen. Het blijkt dan soms lastig om cliënten over een drempel te krijgen om aan het werk te gaan. Er kan wel gebruik gemaakt worden van ontwikkelplaatsen en dagbesteding, maar bij eigen regie is het lastig om mensen op zo'n plek te krijgen als ze niet willen. Het is dan belangrijk dat de klantmanager toch de opties benoemt en probeert de mentale problematiek op te pakken in de begeleiding.

Organisatie

Vooraf tijdens het derde en vierde focusgroepgesprek is ingegaan op de context van de organisatie waarin het experiment zich afspeelt. Er is besproken in hoeverre bijvoorbeeld het uitbesteden van de begeleiding naar Lucrato invloed heeft op de uitkomsten. Maar ook is gesproken over de focus van de gemeente op voltijdse uitstroom naar werk en de integratie van aanwezige budgetten (P-wet, WMO) in de gemeente. Allereerst is er tijdens het derde gesprek besproken of psychosociale begeleiding zou helpen bij een groep waarin gezondheidsproblemen een grote rol spelen. De klantmanagers geven aan dat er nu een programma bestaat waarin mensen een soort van job coaching krijgen en waarbinnen er veel kennis is over verslaving, psychiatrie en combinaties daarvan. Voor dat programma is wel een WMO-indicatie nodig. Het is een combinatie van wat een trajectregisseur doet samen met een andere hulpverlenende partij. Dit zou een model kunnen zijn om op door te denken en wat op veel verschillende manieren zowel intern of extern georganiseerd zou kunnen worden, zolang er maar veel kennis is van de problematiek. Vooralsnog zijn volgens de klantmanagers de WMO-budgetten en de P-budgetten gescheiden. Volgens hen zou het kunnen helpen als er bijvoorbeeld een deel van het P-budget voor dit soort trajecten ingezet zou kunnen worden.

In het vierde focusgroepgesprek wordt opnieuw gesproken over het bundelen van kennis over psychische problematiek binnen de organisatie. Zou het helpen als daar meer kennis over is om het gesprek op gang te krijgen of mensen in beweging te krijgen? De klantmanagers zijn hierover verdeeld. De een vreest voor meer etiketjes plakken, terwijl de ander er wel een meerwaarde in ziet. Het kan er bijvoorbeeld voor zorgen dat cliënten die zich nog niet bewust zijn van hun problematiek, maar wel aanwezig is, op een andere manier benaderd en meer begrepen kunnen worden. De klantmanagers hebben wel voorlichting gehad over hoe ze psychische problemen kunnen herkennen, maar op dit moment is er nog geen vraagbaak aanwezig binnen de organisatie. Zo'n vraagbaak zou op verschillende manieren gefaciliteerd kunnen worden, bijvoorbeeld door een zorgorganisatie erbij te betrekken of door intern mensen beter op te leiden. De klantmanagers zien er wel de meerwaarde van in om kennis gemakkelijk beschikbaar te hebben en met iemand te kunnen sparren als ze er zelf niet uitkomen. Op dit moment is er binnen de organisatie geen prioriteit voor het investeren in gesprekstechnieken, methodieken of specialisatie omdat er nu geen prioriteit wordt gegeven aan re-integratie en participatie, maar aan schuldhelpverlening. Toch blijft een moeilijkheid hierbij dat de gemeente volgens de klantmanagers nog steeds gefocust is op uitstroom en dat dat ook de taakstelling is van de klantmanager, ondanks dat ze redelijk zijn vrijgespeeld in het experiment. Activeringscoaches zouden die druk minder voelen volgens de klantmanagers. Volgens

de klantmanagers moet de gemeente realiteit scheppen in het bestand en inzien dat deze mensen niet direct bemiddelbaar zijn. De begeleidingsvraag is veel groter dan nu wordt gedacht en de uitstroomcijfers zijn daarom niet reëel. Vanuit de politiek wordt het roer te vaak omgegooid en wordt er aangestuurd vanuit de 'korte klap', maar dat is bij deze mensen niet van toepassing. De klantmanagers geven ook aan dat hoe succesvol het aansturen van deze groep is niet zozeer te maken heeft met of de begeleiding wordt uitbesteed aan een externe partij of niet, maar dat het ligt aan de opdracht die de gemeente aan die partij geeft. De opdracht die nu wordt gegeven is gekoppeld aan de financieringsstructuur van de gemeente en die wordt bepaald door het Rijk. De gemeente moet zelf het tekort in de buig-middelen opvangen, waardoor de focus van de gemeente ligt bij uitstroom naar werk, om het buigtekort zo laag mogelijk te houden.

Volgens de klantmanagers is het erg belangrijk dat er wordt erkend dat ongeveer 60% van de mensen in de bijstand een medische of psychosociale problematiek heeft wat vraagt om een andere manier van begeleiden. Er wordt nu vanuit de organisatie wel meer gedaan aan preventie bij de poort. Aan de voorkant wordt dan formeel het gesprek aangegaan richting werk. En dat werkt goed, want er is 40% preventie, maar de uitstroom loopt achter. Maar bij de uitstroom wordt dan geen rekening gehouden met dat er minder mensen instromen en ook niet met uitstroom naar parttime werk. Ook heeft er binnen de gemeente een reorganisatie van werkzaamheden plaatsgevonden, waardoor schuldhulpverlening en trajectregie bij elkaar zijn gekomen. Hierdoor ligt nu de focus op schuldhulpverlening en niet op doorstroom naar Lucrato. De doelmatigheid heeft geen prioriteit meer. Dit betekent voor mensen die voorheen alleen trajectregie deden nu geschoold moeten worden voor schuldhulpverlening. Dat blijkt arbeidsintensief en voor sommige mensen lastig om onder de knie te krijgen. Daarnaast blijkt dat het binnen de groep intensieve begeleiding in de praktijk lastig is om intensief met zowel de schuldhulpverlening als de extra begeleiding bezig te zijn.

In het vierde focusgroepgesprek is er dieper in gegaan hoe deze reorganisatie in de dienstverlening heeft doorgewerkt in het experiment. Vanaf mei 2018 zijn de klantmanagers actief met de schuldhulpverlening aan de slag gegaan en nu ze er meer kennis van hebben, zijn ze er ook steeds drukker mee. Ze krijgen nu veel meer cliënten met schuldenproblematiek. Tijdens de begeleiding is er nu zoveel tijd nodig voor schuldhulpverlening, waardoor er nauwelijks tijd is voor re-integratie. Ook is er minder tijd om andere cliënten te begeleiden, die zijn niet uitgenodigd of zijn minder vaak gesproken (binnen intensieve begeleiding). Impliciet wordt er tijdens de schuldhulpverlening wel gewerkt aan het wegnemen van barrières om aan het werk te gaan, omdat schulden vaak een obstakel zijn om verder te komen.

Beleidsaanbevelingen

Een klantmanager geeft aan dat hij qua sturingsmogelijkheden het soms lastig vindt om oudere mensen aan het werk te krijgen. Voorheen was er vanaf 57 jaar geen sollicitatieplicht, maar dat is komen te vervallen met de Participatiewet. Andere klantmanagers geven aan dat zij juist vaak oudere cliënten hebben die nog wel aan het werk willen en kunnen. Er wordt door de klantmanagers ook aangegeven dat ze vinden dat vrijwilligerswerk begrensd moet worden. In sommige sectoren, bijvoorbeeld de zorgsector zou vrijwilligerswerk geschrapt moeten worden, omdat de krimp in die sector voornamelijk wordt opgevuld met mensen uit de bijstand onder de noemer van vrijwilligerswerk. Op deze manier bouwen

mensen geen pensioen op en vallen ze na hun 67^{ste} jaar in eeuwige armoede. Dat hebben de cliënten soms ook zelf niet goed op het netvlies. Hoewel vrijwilligerswerk maatschappelijk gezien van grote waarde is, verschuiven de grenzen en de vrijblijvendheid van het vrijwilligerswerk over de tijd en wordt het steeds afhankelijker van de persoon die het uitvoert, maar er niet voor wordt betaald en ook geen uitzicht heeft op een betaalde baan. Een oplossing zou volgens de klantmanagers kunnen zijn dat mensen niet langer dan 6 maanden bepaald vrijwilligerswerk doen, zodat het flexibel blijft.

Een andere oplossing voor dit probleem zouden de sociale werkbedrijven kunnen zijn, zodat mensen die niet in de reguliere arbeidsmarkt aan de slag kunnen, wel werk en loon hebben. Volgens de klantmanagers is een nadeel dat dat op basis gaat van loonwaardebepaling, waarbij dit alleen werkt voor mensen die een lage loonwaarde hebben maar niet voor mensen die als ze eenmaal aan de slag gaan, een hogere loonwaarde blijken te hebben.

3.3.4 Samenvatting

Tot slot volgt een korte samenvatting waarin een aantal belangrijke punten uit de procesevaluatie worden benoemd gerangschikt naar onderwerp.

Effect van de treatments:

- Het effect van elke treatment is afhankelijk van de persoon in begeleiding. Begeleiding is maatwerk en elke klantmanager doet het op een andere manier.
- Een andere bejegening van de cliënt, het inspelen op de intrinsieke motivatie en het opbouwen van vertrouwen, zijn de belangrijkste factoren voor een succesvolle begeleiding. Daarnaast is ook het voeren van het bredere gesprek (niet alleen werk) van groot belang.
- Het opbouwen van een vertrouwensband vergt tijd, maar zorgt er wel voor dat de begeleiding soepeler verloopt.
- Er is een aantal barrières die een rol kan spelen in het begeleiden van de cliënt naar werk, zoals langdurig in de bijstand zitten, reisafstanden, ziekte, schulden of verslaving, commitment aan vrijwilligerswerk, het hebben van kinderen onder de 18, niet-gesignaleerde psychische problematieken en nooit hebben gewerkt. Maatwerk is hierdoor onmisbaar alsook tijd en aandacht om de problematieken te doorgronden.

Training en uitvoering:

- De gevolgde trainingen en intervisies waren zinvol en waardevol. Wel is er behoefte aan een duidelijk handelingskader per treatment en regelmatige herhaling van de geleerde methodieken.
- Bij de uitvoering van de treatments liepen klantmanagers soms tegen problemen aan. Bij eigen regie is het moeilijk om niet te sturen en onduidelijk wat er moet gebeuren als de cliënt zelf geen contact opneemt.
- De intensieve begeleiding is erg intensief voor zowel klantmanager als cliënt. Wel is het mogelijk te werken aan kleinere subdoelen die de drempel naar werk kunnen verlagen. In combinatie met de schuldhelpverlening ervaren klantmanagers deze vorm van begeleiding als erg intensief en blijft er in de praktijk minder tijd over per cliënt en ligt de focus niet meer op re-integratie. Aan het eind van het experiment blijft een groep over die weinig beweegt, waardoor het lastig is om de positieve 'flow' of spirit erin te houden.

- Soms missen klantmanagers een stok achter de deur als een cliënt geen initiatief neemt. Soms zou een steuntje in de rug hen wellicht kunnen helpen bepaalde drempels over te gaan.
- In sommige gevallen zou het fijn zijn als klantmanagers na twee jaar een cliënt kunnen overdragen naar een collega, zodat de cliënt kans heeft op een ander gesprek.

Organisatie:

- Omdat een aanzienlijk deel van de bijstandsgerechtigden kampt met een psychische of lichamelijke probleem, zou het een optie kunnen zijn meer kennis over deze problemen binnen de organisatie te ontwikkelen, zodat er altijd een 'vraagbaak' beschikbaar is voor de klantmanagers en zij eventueel sneller kunnen doorverwijzen.
- De organisatie moet meer inzicht krijgen in welke cliënten een registratie hebben in het Doelgroepen Register om zo beter te kunnen inspelen op de behoeften van de cliënt
- Een te grote case-load zorgt ervoor dat cliënten te weinig worden gezien en minder stappen kunnen zetten. Een caseload tussen de 75 en 90 zou realistisch en behapbaar zijn.
- Vanwege de reorganisatie is schuldhulpverlening en trajectregie samengebracht en dit zorgt ervoor dat de nadruk in de begeleiding ligt op de schuldhulpverlening en niet op re-integratie.

Leereffecten klantmanager:

- Klantmanagers geven aan dat ze meer spiegelen met de deelnemer over wat er moet gebeuren.
- Het is fijn om de actie bij de cliënt neer te kunnen leggen en minder te sturen.
- Het aanleren van eigen regie vergt in eerste instantie intensieve aandacht en veel tijd maar later wordt dat minder
- De tijd en aandacht van klantmanagers moet juist gestoken worden in mensen die langer dan 5 jaar in de bijstand zitten en niet in de mensen die aan de voorkant al plaatsbaar zijn. Daarvoor moet de organisatie ook de doelen aanpassen en ervoor zorgen dat de doelen haalbaar zijn, zodat ook de klantmanagers gemotiveerd blijven.

Aanbevelingen voor de toekomstige werkwijze en rol klantmanager:

- Aandacht voor de afbouw van toeslagen vooral in verband met deeltijd, om zo de barrière weg te nemen om meer uren te gaan werken, voornamelijk voor alleenstaande moeders.
- Vrijwilligerswerk moet begrensd worden, om zo te voorkomen dat vrijwilligersorganisaties afhankelijk worden van een persoon. Ook omdat in sommige krimpsectoren (zoals de zorg) deze krimp wordt opgevangen met vrijwilligerswerk; vrijwilligerswerk niet altijd uitzicht biedt op een betaalde baan, maar mensen er wel aan hechten en omdat vrijwilligers geen pensioen opbouwen en dus later in armoede vervallen.
- Sociale werkbedrijven zouden een oplossing kunnen zijn voor mensen die niet op de reguliere arbeidsmarkt aan de slag kunnen. Het nadeel is wel dat dit gebaseerd is op een loonwaardebepaling en alleen werkt voor mensen die een blijvend lage loonwaarde hebben.

4 Bevindingen en conclusies

4.1. Opzet van het experiment en werving

De start

De gemeente Apeldoorn/Epe was samen met Apeldoorn in 2017 nu ruim 2 jaar geleden een van de voortrekkers in Nederland om een experiment met regelluwe bijstand te starten in de Participatiewet buiten de kaders van de AMvB om. Niet-AMvB gemeenten hebben meer vrijheidsgraden bij het ontwerp van het experiment dan de AMvB gemeenten doordat zij niet gebonden zijn aan de nogal als beperkend ervaren voorwaarden van artikel 83. Anderzijds hebben zij niet de mogelijkheid om de vrijlatingsbepalingen te verruimen. De gemeente wilde een nieuwe aanpak waarin wederzijds vertrouwen tussen cliënt en gemeente alsook eigen regie en maatwerk belangrijke uitgangspunten zijn van de begeleiding. Daarom wordt het experiment ook wel “Zelf aan het stuur en op maat” genoemd. Bij de opzet van het experiment is overigens aangesloten bij de opzet van de experimenten in de AMvB gemeenten. Dat betekent dat ook in Apeldoorn/Epe en de andere niet-AMvB gemeenten is gekozen voor een ‘random control trial’ (RCT) design waarbij deelnemers aselekt worden toegewezen aan de ‘treatments’ of behandelingsgroepen (zie onderzoeksvoorstel: Muffels & van der Klink, 2017). De treatment groepen zijn wel verschillend doordat niet-AMvB gemeenten niet mogen experimenteren met een ruimere vrijlating maar dus wel met versoepeling van de verplichtingen (eigen regie) en een andere bejegening en begeleiding van de cliënt (maatwerk en intensieve begeleiding).

Werving en onderzoek: invulling van de vragenlijsten

De werving zou voor de helft uit het zittende bestand en voor de helft uit nieuwe instroom plaatsvinden. Om dat voor elkaar te krijgen werd er een peildatum bepaald voor deelnemers met een lopende uitkering. De berekende peildatum werd 1 juli 2016. Voor de werving van de nieuwe instroom was 6 tot 9 maanden ingeruimd, tussen 1 oktober 2017 en 1 augustus 2018. Er is veel aan gedaan om de werving tot een succes te maken met voorlichtingsmateriaal, filmpjes, voorlichtingsbijeenkomsten etc. De werving duurde langer dan gedacht omdat het lastig bleek nieuwe instromers te motiveren mee te doen. De laatste deelnemer stroomde op 25 september 2018 binnen. Uiteindelijk leidde dit tot 787 aanmeldingen waarvan 199 aanmeldingen niet voldeden aan de toelatingscriteria. Daarmee kwam het aantal toegelaten aanmeldingen op 588. Na aanmelding en het eerste gesprek met de consulent waarin de loting werd bekend gemaakt konden mensen zich alsnog afmelden. In Apeldoorn meldden 87 mensen zich alsnog af en in Epe 9. Daarmee werd een netto aantal deelnemers bereikt van 411 deelnemers in Apeldoorn en 81 in Epe. Voor de analyse nemen we echter ook de afmeldingen mee omdat zij in beginsel gestart zijn met de treatment maar deze voortijdig afbraken. Dit noemen we de ‘intention to treat’ (ITT) analyse. Het streefaantal van 540 (450 in Apeldoorn en 90 in Epe) werd in de periode tussen 1 oktober 2017 en 1 oktober 2018 dus gehaald. De invulling van de Qualtrics online vragenlijsten door de deelnemers bleek evenwel in de fase daarna geen sinecure. Niettemin is een respons van 67% gehaald, 93% in de nulmeting en 53% in de vervolgmetingen. Door de gunstige arbeidsmarkt nam de uitstroom naar betaald werk toe van 15% gemiddeld naar 20% waardoor het aantal deelnemers in begeleiding afnam maar tegelijkertijd ook de respons werd beïnvloed. Daarnaast stroomden nog eens 15% van de deelnemers uit vanwege verhuizing, herziening recht op uitkering en dergelijke. Eenmaal uitgestroomd werd de motivatie minder

om de vragenlijst in te vullen. De klantmanagers vulden tijdens het proces van uitvoering ook een vragenlijst in en deze is voor meer dan 80% tijdens het experiment drie keer ingevuld voor alle deelnemers. De gunstige arbeidsmarkt was plezierig voor de deelnemers zelf, een op de vijf deelnemers vindt werk, maar daardoor liep het experiment sneller leeg dan verwacht. Het gevolg was dat de kansrijkste deelnemers naar werk uitstromen en de klantmanagers te maken kregen met deelnemers die moeilijker aan het werk kwamen. Dat maakte de begeleiding c.q. coaching na meting 2 vanuit het perspectief van de klantmanager een lastiger opgave zoals ook uit de focusgroep gesprekken met de klantmanagers/consulenten volgde.

4.2 Uitvoering van de treatments

Het experiment beoogt voor deze deelnemers een alternatieve aanpak en bejegening te ontwikkelen uitgaande van vertrouwen, eigen regie, keuzevrijheid, zelfredzaamheid en maatwerk. Niet de uitkomst staat centraal maar het proces van begeleiding en ondersteuning waarin mensen wordt geleerd zelfredzaam te worden, dus zelf aan het stuur en afgestemd op de noden en mogelijkheden van de deelnemer (vraaggericht maatwerk). Deze aanpak is anders en wordt door de klantmanagers als verfrissend, en vernieuwend ervaren. Het werk als professional werd er boeiender door, relaxter ook, en men ging met meer plezier naar het werk. Het 'echte' gesprek, aandacht op basis van vertrouwen en de eigen regie bij de deelnemer werden als de belangrijkste 'triggers' van de andere aanpak gezien en ervaren. Het werd gezien als een uitdaging. Die uitdaging betreft de treatment van 'eigen regie' maar ook de treatment van intensieve begeleiding waarbij de klantmanagers/consulenten een andere aanpak voorstaan en de filosofie erop gericht was om mensen zelfredzaam te maken en de deelnemer te leren de eigen verantwoordelijkheid te nemen en die ook waar te maken. Daartoe hebben de klantmanagers/consulenten op basis van de 'skills' methodiek een training gekregen van Roland Blonk waarin vaardigheden zijn geleerd om de zelfredzaamheid van de deelnemers te bevorderen. Eigen regie is dus geen 'loslaten' maar 'leren zelfredzaam te zijn of te worden' door het gesprek aan te gaan, het initiatief bij de deelnemer te laten en handelingsvrijheid te geven om zelf in actie te komen. Overigens vond men het onderscheid zeker in het begin tussen eigen regie en intensieve begeleiding moeilijk te duiden en te implementeren. Men had behoefte aan praktische handvatten die er in hun visie onvoldoende waren. Werkende weg werd middels de 'skills' training en de focusgroep gesprekken en onderlinge intervisiebijeenkomsten een praktische oplossing gevonden.

De deelnemers in Apeldoorn/Epe worden na de reorganisatie per 1 januari 2018 deels begeleid door klantmanagers die zijn aangesteld bij het werkbedrijf Lucrato en anderzijds door consulenten werkzaam bij de afdeling Activering en Inkomen. De klantmanagers van het Werkbedrijf Lucrato begeleidden mensen met een korte afstand tot de arbeidsmarkt (korter dan 1 jaar) en de consulenten van A&I deelnemers die langer dan 1 jaar in de uitkering zitten en lager op de participatieladder staan, dus minder snel kans hebben op een betaalde baan. Drie maanden nadat het experiment startte op 1 oktober 2017 was dus net een reorganisatie van het gehele ondersteuningsproces gestart. Het motto van de reorganisatie was "Iedereen doet mee" en naar een meer 'mensgerichte' in plaats van 'systeem' of 'wet' gerichte aanpak (Beijer, 2017). Dat laatste betekende bijvoorbeeld dat er meer aandacht is voor een vraaggerichte aanpak (wat wil de bijstandsgerechtigde?) en individueel maatwerk. Daardoor veranderde

evenwel de reguliere dienstverlening en ging deze wat al meer lijken op de treatments in het experiment waarin vraaggerichtheid en maatwerk ook sleutelbegrippen waren. Vanwege het experiment werden de klantmanagers van Lucrato en de consultants van Activering en Inkomen bij elkaar gebracht in het team Vertrouwensexperiment dat onder begeleiding van twee projectleiders Janneke Oude-Alink en Lieke Munsters aan de slag ging om het experiment uit te voeren. Naast de training werden in groepsbijeenkomsten moeilijke casuïstiek besproken die vervolgens ook aan de orde werden gesteld in het focusgroepgesprek. 'Al doende leert men' en men vond naar eigen zeggen geleidelijk aan een weg om de treatments zo goed mogelijk uit te voeren. De bijeenkomsten gaven ook de mogelijkheid van intervisie, het bij de les houden van elkaar voor de aanpak waarvoor was gekozen in het experiment.

Eerste bevindingen van de procesevaluatie

In de procesevaluatie (de drie door de consultants ingevulde vragenlijsten per deelnemer) vinden we treatment effecten maar in een andere richting dan we hadden verwacht. De standaardgroep doet het volgens de consultants althans beter op veel uitkomstmaten dan de beide andere treatments eigen regie en intensieve begeleiding. Uit de ingevulde vragenlijsten blijkt dat de consultants weinig vertrouwen hebben in de 'andere aanpak' en de standaard aanpak geschikter vinden. De traditie in Apeldoorn kenmerkt zich door een 'stevige controlerende' aanpak. Dat beeld werd bevestigd in de focusgroep gesprekken maar in mindere mate. Men toonde zich in die gesprekken enthousiast voor het gedachtegoed van het experiment, men steunde het idee van vraaggericht maatwerk, tijd voor het 'echte' gesprek en ook het idee van zelfredzaamheid maar wel met een stevige vinger aan de pols. Gevraagd naar of de eigen treatment geschikt is of een andere treatment beter voor deze deelnemer werd veelal verwezen naar de standaard treatment als een betere optie. Zelf heeft men ook een voorkeur voor de standaard treatment. Die treatment kent men, daar weet men zijn of haar weg in te vinden en men acht die effectiever. Toch geeft men het experiment gemiddeld een 8 qua gedachtegoed maar men heeft wel vraagtekens of het is uitgevoerd conform dat gedachtegoed. Men mist handvatten en een methodiek om de treatments uit te voeren zoals bedoeld en op een voldoende onderscheidende manier. Vooral de extra tijd die men krijgt voor het 'goede' gesprek lijkt te worden gewaardeerd. Dat neemt niet weg dat er in de procesevaluatie ook veel positieve aanknopingspunten te vinden zijn bijvoorbeeld doordat op veel uitkomstmaten zoals de motivatie voor het hoogst bereikbare doel, realistische eisen stellen, zelf in actie komen, zelf doelen stellen, stijging op participatieladder en leereffecten bij de consultant er vooruitgang in de tijd valt waar te nemen op basis van de ingevulde vragenlijsten door de consultants. Zeker ook de stijging op de participatieladder over tijd is een belangrijke constatering die door de gemeente wordt gezien als een adequaat instrument om de vooruitgang bij de bijstandsgerechtigde in termen van sociale participatie of sociale stijging te peilen en te monitoren. Opvallend is dat de consultants weinig vertrouwen lijken te hebben in het effect van de treatments op de deelnemer. Het verwachte effect van de treatment voor de deelnemer werd als negatief ingeschat evenals de leereffecten bij de deelnemer. De andere effecten zijn over tijd allemaal positief en significant voor alle treatments inclusief de standaard treatment. De intensieve begeleidingsgroep doet het op een aantal aspecten volgens de consultants beter dan de eigen regiegroep, zoals motivatie voor eigen treatment, leereffect van de deelnemer, sociale participatie en stijging op de participatieladder en op zijn minst net zo goed als de standaard treatment op alle aspecten.

Voorts wordt uit de bevindingen van de procesevaluatie klip en klaar dat de standaard treatment niet hetzelfde is gebleven als die voorheen was en dat deze is veranderd en in belangrijke mate in de richting van de andere treatments is opgeschoven waardoor we bijvoorbeeld geen verschillen vinden voor de treatment intensieve begeleiding vergeleken met de standaardgroep en negatieve voor de eigen regie treatment. Daardoor is de standaardgroep mogelijk geen goede vergelijkingsmaatstaf. De vooruitgang in de tijd voor alle groepen doet vermoeden dat de andere aanpak succesvol is geweest op de verschillende uitkomstmaten, zowel op het vlak van motivatie, gedrag en participatie uitkomsten.

Eerste bevindingen productevaluatie

Of de treatments zijn uitgevoerd conform het beoogde design is vooralsnog moeilijk te zeggen. Daarvoor zijn nadere analyses nodig van de inhoud van de begeleiding en ondersteuning en de contacten die er zijn geweest met de deelnemer. Deze analyse ontbreekt nog. Gelet op de uitkomsten van de focusgroep gesprekken en de analyse van de consultant vragenlijsten zijn we niet bijzonder optimistisch over de kans dat er geen sprake ook is van 'implementation failures' bij de uitvoering van het experiment ondanks de goede intenties. Er zijn zeker ook 'theory failures' aangezien de theorie weinig handvatten bood voor een praktische vertaling van de treatments in de uitvoering. Wat ontbreekt is een goede vertaling van de kennistheoretische onderbouwing van de treatments in een werkzame methodiek.

De analyse van de drie deelnemersvragenlijsten laat significante verschillen zien op enkele uitkomstmaten zoals zelfredzaamheid en armoede maar op vele andere uitkomstmaten zoals keuzemogelijkheden, sociaal vertrouwen, sociale participatie en sociale netwerken geen verschillen. In de tijd vinden we positieve en significante effecten voor mentale gezondheid en welbevinden voor alle groepen. We vinden geen verschillen tussen eigen regie en intensieve begeleiding. Veelal zijn de gevonden geringe aantal significante effecten in vergelijking met de standaardgroep sterker voor eigen regie. Deze resultaten zijn volgens ons relevante en belangrijke bevindingen, mede omdat de vergelijking met de standaardgroep problematisch is om eerder genoemde redenen. Zouden we een betere vergelijkingsgroep hebben zoals bij de uitstroom naar werk dan verwachten we andere resultaten.

Verandering tijdens de uitvoering

De wereld staat tijdens het experiment evenwel niet stil. De politieke en beleidscontext verandert maar er zijn ook leereffecten bij de coaches waardoor de treatments ook veranderen. Dat maakt het experiment complex want voor een juiste vergelijking zouden de treatments zoveel mogelijk hetzelfde moeten blijven. De verandering in uitvoeringspraktijk en beleidscontext, zoals de reorganisatie in Apeldoorn in 2018, beïnvloedt de standaard treatment waardoor een zuivere vergelijking tussen de reguliere en alternatieve treatments in het onderzoek werd bemoeilijkt. Het eindverslag gaat daar op verschillende plaatsen nader op in. Uit de proces- en productanalyses wordt duidelijk dat de standaard treatment is mee geëvolueerd met de andere treatments. De lijntjes lopen in de tijd bij bijna alle drie treatment groepen en de meeste uitkomstmaten zowel in de product als de procesevaluatie min of meer gelijk op. We vinden significante treatment effecten in de procesevaluatie in de vergelijking tussen treatment groepen maar in een andere richting dan we verwachtten en we vinden significante treatment effecten in de productevaluatie (analyse deelnemersvragenlijsten) bij bepaalde uitkomsten zoals zelfredzaamheid en armoede. Bovenal vinden we significante effecten over tijd voor alle treatment groepen zoals voor welbevinden en mentale

gezondheid. De standaard treatment is veranderd waardoor we geen effecten kunnen vinden in de vergelijking met de standaardgroep. Dat is deels ook vanzelfsprekend omdat beleid mee evolueert met de context en de inhoud van de standaard aanpak over de tijd maar deels heeft het ook te maken met een soort “Hawthorne” effect, een effect van het experiment en het onderzoek waardoor mensen zich anders gaan gedragen zowel de deelnemer als de consultant.

4.3 Uitstroom naar werk en matching

Bij de meting van de uitkomstmaat volledige uitstroom naar werk vonden we significante verschillen tussen de beide treatments met de standaardgroep en de referentiegroep. De verschillen met de referentiegroep waren veel groter dan met de standaardgroep. Deze laatste waren in de orde van grootte van 5 tot 7% voor Apeldoorn voor de vergelijking met de standaardgroep en 10 tot 13% in vergelijking met de referentiegroep. Er is dus een groot verschil tussen de vergelijking met de standaardgroep en de referentiegroep. Het vermoeden bestaat mede op grond van de bevindingen in de procesevaluatie dat de aanpak in de standaardgroep ook is veranderd tijdens het experiment. Voor de ‘uitstroom naar werk’ uitkomstmaat bleek dat probleem te kunnen worden opgelost door een vergelijking van de uitstroom niet alleen tussen de treatments en de standaardgroep maar met name met de referentiegroep, de groep van bijstandsgerechtigden die niet aan het experiment heeft deelgenomen maar wel tot dezelfde doelpopulatie behoort en daarbij te corrigeren voor verschillen in de samenstelling tussen deelnemers en niet-deelnemers die is ontstaan door zelfselectie bij aanmelding. Voorts zijn gegevens in de vragenlijst gebruikt over de mate waarin men in meting 2 en 3 aan het werk is gegaan in een deeltijd of een voltijd baan. Tevens is gevraagd of men vrijwilligerswerk heeft verricht. Tot slot is een vraag gesteld over het soort contract waarin men werkt als men werk heeft gevonden; is dat tijdelijk of vast werk of werk als zelfstandige. Op basis van deze gegevens bleek bijvoorbeeld dat 14% van de deelnemers zegt in deeltijd te werken en 20% zegt werk te hebben gevonden met een tijdelijk of vast contract, waarvan twee-derde een tijdelijk contract is. Vergelijken we de drie groepen dan zien we dat bijna 25% van de deelnemers in de intensieve begeleidingsgroep werk heeft gevonden tegenover 16% in de controlegroep. Het verschil zit niet in de mate waarin men in deeltijd gaat werken, dat was in beide groepen 14-16%, maar in het aantal voltijd banen. Deze cijfers geven echter geen volledig beeld aangezien slechts ongeveer 55% van de deelnemers de tweede en derde vragenlijst heeft ingevuld.

Omdat de referentiegroep niet is gerandomiseerd weten we niet zeker of de grotere uitstroom in de treatments ten opzichte van deze groep een causaal effect is. Daarnaast heeft er zelfselectie plaatsgevonden bij de aanmeldingen. Zo bleek dat in de onderzoekpopulatie korte dueren waren oververtegenwoordigd en lange dueren ondervertegenwoordigd. Deels had dat ook te maken met het feit dat afmeldingen met name plaatsvonden bij mensen uit het zittende bestand met langere dueren. Ook naar nationaliteit was de groep niet representatief. Om hiervoor te corrigeren is ‘matching’ toegepast. Daarbij zijn er twee matching methoden gebruikt: ‘nearest neighbour’ (wegingsmethode voor de onderliggende variabelen) en ‘propensity score’ matching (gebaseerd op verklarend kans model). Het idee was om uit de referentiegroep een groep van 564 mensen (inclusief afmeldingen) te vinden die nagenoeg qua achtergrondkenmerken 100% overlap vertonen met de deelnemers. Op die manier kon worden gecorrigeerd voor de selectiviteit die was ontstaan bij de deelnemers groep ten opzichte van de

bijstandspopulatie. Dat bleek heel goed mogelijk, de kwaliteit van de matching was goed, en leidde ertoe dat het gevonden treatment effect naar schatting geen 5-13% (uitstroom per jaar) bedraagt maar 4 tot 7.5% na correctie voor selectie, maar nog altijd substantieel.

4.4 Niet-werk gerelateerde uitkomstmaten

De meeste treatment effecten die we vonden bij de niet-werk gerelateerde uitkomstmaten bleken niet significant in vergelijking tussen de beide treatment groepen en de standaardgroep en we vonden ook geen significante verschillen tussen eigen regie en de standaardgroep. De meeste hypothesen die we formuleerden aan het begin worden daarom niet bevestigd met uitzondering van de hypothesen ten aanzien van zelfredzaamheid en armoede. Daar vinden we positieve en significante treatment effecten voor eigen regie en intensieve begeleiding vergeleken met de standaardgroep in de productevaluatie in de metingen bij de deelnemers. We vonden geen treatmenteffecten bij welbevinden, gezondheid, keuzemogelijkheden en vertrouwen. We vonden duidelijke aanwijzingen in de procesevaluatie maar ook in de uitstroom naar werk analyse voor een verandering van de treatment in de standaardgroep. Zelfs in de analyse van de deelnemers vragenlijsten zijn daarvoor aanwijzingen te vinden. Bij de uitkomstmaat van welbevinden (niet de minst belangrijke in de literatuur) vonden we een positief significant treatment effect over de tijd bij de standaardgroep. Indien de standaard treatment niet was veranderd zou het vinden van een dergelijk effect onwaarschijnlijk zijn geweest. Met andere woorden, indien de standaard treatment inderdaad is veranderd, is geen zuivere vergelijking met de treatment groepen meer mogelijk. Dan zullen we een vergelijking moeten maken met een vergelijkingsgroep die niet beïnvloed is door het experiment. Op dit moment is zo'n alternatieve vergelijkingsgroep niet beschikbaar.

Effecten over tijd bij de andere uitkomstmaten

We vinden echter wel positieve en significante effecten over tijd in verband met welbevinden en mentale gezondheid maar dan voor alle groepen. Of dit het gevolg is van het experiment, dus een treatment effect, is vooralsnog causaal niet vast te stellen zonder zo'n vergelijkingsgroep. Het is wel opvallend omdat er geen evidentie is dat welbevinden bij bijstandsgerechtigden zou zijn toegenomen in de afgelopen twee jaar, integendeel cijfers van het SCP (Wennekers e.a. 2019) suggereren dat voor mensen met een score op tevredenheid met het leven of geluk van 6 of lager (waartoe ook de bijstandsgerechtigden behoren) de score de afgelopen jaren – dus na de economische crisisjaren - is verslechterd. Nader onderzoek moet uitwijzen of onderliggend de stijging in verband met welbevinden en mentale gezondheid niet toch een treatment effect is dat niet zichtbaar wordt omdat de standaard treatment gedurende het experiment is gaan lijken op de andere treatments.

Verschillen tussen de treatments

Opvallend is dat we geen verschillen vinden tussen de treatments. Dit heeft verschillende oorzaken. De eerste is dat in het experiment met opzet is gekozen voor een ontheffings -of 'eigen regie' groep die mensen leert zelfredzaam te worden. Dat betekent dus geen 'loslaten' zoals stricto sensu onder het begrip 'eigen regie' zou kunnen worden verstaan. De training aan de klantmanagers/consulenten was gericht op leren zelfredzaamheid waarbij het onderscheid tussen eigen regie en intensieve begeleiding minder aandacht heeft gekregen. De treatments zijn aan het begin van het experiment nader omschreven en op schrift gesteld. Daarin werd ook aangegeven hoeveel contactmomenten er gemiddeld over de tijd zullen

zijn per jaar met de klant. Het was duidelijk dat bij eigen regie het aantal contactmomenten minder zou zijn en dat de regie voor contact bij de deelnemer zou behoren te liggen. De afgesproken 'case load' was voor eigen regie ook hoger vastgesteld waardoor er minder tijd zou zijn dan bij intensieve begeleiding per deelnemer. De ervaring leerde echter al snel dat de op papier gestelde uitgangspunten in de praktijk anders uitpakten. Reeds bij het begin van het experiment bleek dat 'leren zelfredzaamheid' in het begin juist veel contact vraagt met de deelnemer en dat de meeste deelnemers het juist prettig vonden in het experiment dat ze veel contact hadden met de klantmanager. De gangbare praktijk was dat mensen die reeds enkele jaren in de bijstand zaten, het zogenaamde 'granieten bestand', vaak heel weinig of geen enkel contact meer hadden met de klantmanager. Vaak werden mensen al ontheven van de arbeids – en/of sollicitatieverplichting. Deelname aan het experiment werd door de deelnemer dan ook soms gezien als een manier om meer ondersteuning te krijgen en meer contact. Door anderen werd eigen regie wel geïnterpreteerd als 'loslaten' en handelingsvrijheid. Doordat vanwege de focus op zelfredzaamheid het onderscheid tussen beide treatments niet geheel duidelijk was en ook de contactintensiteit aan het begin niet verschilde was het voor klantmanagers niet van meet af aan duidelijk wat de verschillen waren tussen de treatments. Geleidelijk aan is het onderscheid tussen eigen regie en intensieve begeleiding in de werkwijze van de klantmanager/consulent helderder geworden en later bleek ook dat er minder contact met de deelnemer nodig was. Dit heeft tijd gevraagd. Daarnaast is er andere uitstroom vanwege verhuizing, overgang naar een andere regeling, herzien recht op uitkering, arbeidsongeschiktheid en dergelijke waardoor het aantal deelnemers afneemt. Na 2 jaar was ruim 35 % uitgestroomd. Dit heeft mogelijk effect gehad op de gevonden verschillen in uitkomsten tussen de beide treatments. Ook al werd van hen gevraagd om de treatments zo zuiver mogelijk uit te voeren en niet te veranderen, vanuit moreel en ethisch oogpunt vonden klantmanagers het lastig om de deelnemer iets te onthouden wat op dat moment door de klantmanager als professional nodig wordt geacht. Niet uitgesloten kan worden dat dit een effect heeft gehad op het zuiver uitvoeren van de treatments.

4.5 Tot besluit

We vinden dus positieve resultaten op uitstroom naar werk zowel bij vergelijking met de gerandomiseerde controlegroep als in vergelijking met de niet-deelnemers. Maar bij de andere uitkomstmaten zien we dat enkele van onze verwachtingen op deze andere uitkomstmaten dan werk worden bevestigd en anderen worden weerlegd op basis van deze cijfers. Vooralsnog worden de meeste hypothesen over de verschillen tussen de treatments en de standaardgroep op deze 'zachtere' uitkomstmaten weerlegd. Toch is er ook goed nieuws te melden. De gevonden positieve treatment effecten op zelfredzaamheid en armoede zijn zeker niet onbelangrijk aangezien stimuleren zelfredzaamheid een van de onderliggende doelen was en armoede een serieus probleem vormt zijn dit geen onbelangrijke winstpunten. De scores die klantmanagers geven aan de deelnemers op de participatieladder lieten een significante stijging zien voor de meerderheid van de deelnemers waardoor de afstand tot de arbeidsmarkt werd verkleind. Deze sociale stijging was voor de gemeente ook een belangrijk doel. Eerder constateerden we al dat er positieve en significante effecten zijn op volledige uitstroom naar werk voor de beide treatments maar ook bij deze uitkomstmaat van sociale stijging is dus sprake van een duidelijk treatment effect. We vonden voorts dat bij de intensieve begeleidingsgroep veel mensen in deeltijd te werken, maar niet meer als bij de controlegroep. Wel zijn meer deelnemers van

deze groep vergeleken met de controlegroep in voltijd gaan werken of in een tijdelijk contract. De lijntjes die het verloop over de tijd weergeven van de niet-werk gerelateerde uitkomstmaten laten bijna allemaal een stijging zien ook bij de standaardgroep. Nader onderzoek met een andere vergelijkingsgroep is dus nodig om tot definitieve uitspraken te komen over de uitkomsten op deze andere uitkomstmaten.

Beleidsvragen

Niet alle beleidsvragen die aan het begin zijn geformuleerd kunnen worden beantwoord op dit moment. We laten ze hieronder kort de revue passeren:

- *De eerste beleidsvraag had betrekking op het effect van ontheffing.* Doordat we geen significant verschil vonden tussen de eigen regie en de intensieve begeleidingsgroep is het moeilijk om het effect van ontheffing en intensieve begeleiding afzonderlijk te meten. Voor mensen die in het zogenaamde 'granieten' bestand zitten is in feite al sprake van ontheffing. Er is nog maar sporadisch contact met de (werk)consulent en velen hebben ook ontheffing van de sollicitatie -en arbeidsplicht. Het verschil tussen eigen regie en de referentiegroep van het zittende bestand zou dan een goede vergelijking zijn. Het verschil in uitstroomkansen naar werk tussen beide groepen is ruim 10%. In de ontheffingsgroep meten we niet zozeer het effect van ontheffing als wel het effect van de combinatie van ontheffing en van sturing op zelfredzaamheid. Alle uitkomstmaten wijzen in de goede richting maar zoals eerder vastgesteld is er geen significant effect in vergelijking met de standaardgroep, mogelijk omdat die standaardtreatment zelf is veranderd en meer is gaan lijken op de andere treatments. Voor een deel hangt dat samen met verandering van de beleid -en uitvoeringscontext waarin het gedachtegoed van het experiment geleidelijk aan is geïmplementeerd in de praktijk van de uitvoering (zie hiervoor).
- *De tweede beleidsvraag had betrekking op het effect op de arbeidsinschakeling wanneer mensen op een positieve vraaggerichte manier worden ondersteund en intensiever en op maat worden begeleid (met name intensieve begeleidingsgroep).* Deze onderzoeksvraag is hierboven beantwoord. Er is geen of een klein verschil in uitstroomkansen met de andere treatment groep of met de standaard of controlegroep maar wel met de referentiegroep die volgens ons de juiste vergelijkingsgroep is. De uitstroom is dan fors hoger en na correctie voor selectiviteit nog steeds zo'n 5 tot 13% hoger in de beide treatment groepen vergeleken met de referentiegroep.
- *Een derde beleidsvraag voor de gemeente was of mensen die deelnemen aan het experiment meer ruimte ontwikkelen om eigen initiatief te ontplooiën, meer in brede zin gaan participeren, en wat het effect hiervan is op de arbeidsinschakeling (eigen regie, intensieve begeleiding).* Ook deze vraag kan bevestigend worden beantwoord. Hoewel ook hier evidentie is dat de inhoud van de begeleiding in de standaard/controlegroep ook is veranderd blijft er een significant effect voor beide treatments op zelfredzaamheid en armoede te bestaan. De resultaten van de proces -en productevaluatie tonen dat beide aan. In de procesevaluatie vinden we voorts een positief verschil tussen eigen regie en de intensieve begeleidingsgroep. Op basis van de procesevaluatie denken we dat het werkzame bestanddeel hier lijkt te zijn vertrouwen, aandacht, maatwerk en vraaggerichte ondersteuning.
- *Een belangrijke vraag voor de gemeente was wat de neveneffecten zijn van de ruimere handelingsvrijheid op maatschappelijke participatie, maar ook op zelfredzaamheid, welbevinden en*

gezondheid (eigen regie, intensieve begeleiding). Op zelfredzaamheid zijn de effecten van de procesevaluatie positief ten opzichte van de standaard -of controlegroep. Bij de productevaluatie vinden we ook een significant effect van de beide treatments op zelfredzaamheid alsook op daling van armoede. We vinden voorts significante effecten voor mentale gezondheid en welbevinden over de tijd maar dan voor alle treatment groepen tezamen. We vermoeden dat hier een treatment effect aanwezig is maar daarvoor hebben we een andere vergelijkingsgroep nodig.

- *In het experiment wordt onderzocht wat de uitkomsten over de tijd gezien zijn van meer handelingsvrijheid en een intensievere begeleiding op maat. Dat vraagt een andere manier van ondersteuning en andere kwalificaties van de klantmanagers/consulenten. De gemeente wil graag weten hoe dit proces verloopt. Worden de treatments uitgevoerd zoals beoogd en in welke mate zijn ze onderscheidend? Wat werkt in de uitvoering en wat werkt niet? Welke condities, aspecten maken het verschil en wat betekent dit voor de implementatie van de uitkomsten (procesevaluatie).* Het bleek lastig om de treatments zoals beoogd uit te voeren. Het onderscheid tussen eigen regie en intensieve begeleiding was bij de start niet duidelijk voor de klantmanagers. Of dit een 'theory failure' (theorie fout) of 'implementation failure' (uitvoeringsfout) is, is moeilijk te zeggen. Wij denken beide. In de oratie van Blonk (2018) wordt opgemerkt dat er weinig literatuur is over het integratieve gedragsmodel wat betreft de handvatten die worden aangereikt voor de professional. Wellicht heeft dat een rol gespeeld en is het model nog niet voldoende uitgekristalliseerd in de praktijk. Ook in de uitvoering was er geen methodiek om de treatments inhoud en vorm te geven. Het onderscheidende van de treatments was daardoor minder zeker in het eerste half jaar. We vermoeden dat de verandering bij de standaard treatment (controlegroep) duidt op een 'Hawthorne' of experiment effect waarvoor we niet konden corrigeren. De 'andere bejegening' maakt naar onze visie het verschil hoewel het volledige effect pas na een tijd zichtbaar zal zijn bij de deelnemers zelf. De analyses van de procesevaluatie-vragenlijsten ingevuld door de klantmanagers en de focusgroep gesprekken geven ons de indruk dat het opbouwen van een persoonlijke relatie, vertrouwen geven en aandacht, meer tijd voor het 'echte' gesprek, sturen op intrinsieke motivatie, een luisterend oor en loslaten sleutelbegrippen zijn voor het succes van de aanpak zoals die door de consulenten is ervaren.

Literatuur

- Abbring, J.H., G.J. van den Berg, P. Mullenders en J.C. van Ours (1996). Sancties in de WW; een werkend perspectief. In: Economisch Statistische Berichten, jg. 81, p. 750-753.
- Abbring, J.H., G.J. van den Berg en J.C. van Ours (2005). The Effect of Unemployment Insurance. Sanctions on the Transition Rate from Unemployment to Employment. In: Economic Journal, jg. 115, nr. 505, p. 602-630.
- Atkinson, A. B. (1996) 'The case for a participation income', Political Quarterly, 67, 1, 67–70.
- Bigotta, M., Bonoli, G., Fati, F., Lalive, R., & Oesch, D. (2018). Helping social assistance recipients find jobs: Evidence from a controlled trial (ESPANet working paper, University of Lausanne, Switzerland)
- Bohnet, I., Frey, B. S. and Huck, S. (2001) 'More order with less law: on contract enforcement, trust and crowding', The American Political Science Review, 95, 1, 131–44.

- Bond, G. R., Drake, R. E., & Luciano, A. (2015). Employment and educational outcomes in early intervention programs for early psychosis: A systematic review. *Epidemiology and Psychiatric Sciences*, 24(5), 446–457. <https://doi.org/10.1017/S2045796014000419>
- Blonk, R.W.B., M.W. van, Twuijver, H.A. van de Ven en A.M. Hazelzet (2015). Quicksan wetenschappelijke literatuur gemeentelijke uitvoeringspraktijk. Leiden: TNO.
- Blonk, Roland (2018), We zijn nog maar net begonnen, Openbare rede prof. dr. R. Blonk, Tilburg University op 6 April 2018.
- Card, D., Kluve J., en A. Weber (2010), Active labour market policy evaluations: a meta-analysis, *The Economic Journal*, 120 (November), F452–F477.
- Card, D., Kluve, J., & Weber, A. (2015). What works? A meta-analysis of recent active labour market program evaluations (No. w21431). National Bureau of Economic Research.
- CBS (2017), De Geluk meter, CBS Statline, Voorburg/Heerlen.
- Cuelenaere, B., E. van den Brink en J. Leenen (2019). Ervaringen gemeenten met Participatiewet. Rapportage derde meting. Tilburg: Centerdata.
- Deaton, A., Nancy Cartwright (2017). Understanding and misunderstanding randomized controlled trials, *Social Science & Medicine*, Vol. 210(2018):2-21(<https://doi.org/10.1016/j.socscimed.2017.12.005>)
- Deci, E. L. and Ryan, R. M. (1985) *Intrinsic Motivation and Self-Determination in Human Behavior*, New York: Plenum.
- Delos, S. en J. Oude Alink (2016), Experiment regelluwe bijstandsverlening, Brief aan portefeuillehouders, Werk en Inkomen, Gemeente Apeldoorn/Epe.
- Diener, E., Suh, E. M., Lucas, R. E., & Smith, H. L. (1999). Subjective well-being: Three decade of progress. *Psychological Bulletin*, 125(2), 276–302. <https://doi.org/10.1037/0033-2909.125.2>.
- Dohmen, T., A. Falk, D. Huffman and U. Sunde (2008), Representative trust and reciprocity: Prevalence and determinants, *Economic Inquiry*, 46 (1), 84–90.
- Fehr, E. & Gächter, S. (2000). Fairness and Retaliation: the economics of reciprocity. *The Journal of Economic Perspectives*, 14(3), 159-181.
- Forget, E. L. (2011). The town with no poverty: the health effects of a Canadian guaranteed annual income field experiment. *Canadian Public Policy*, 37(3), 283-305.
- Fehr, E. and Schmidt, K. M. (2003) 'Theories of fairness and reciprocity: evidence and economic applications', *Advances in Economics and Econometrics*, Econometric Society, Eighth World Congress, 1, 208–57.
- Fishbein, M., & Ajzèn, I. (2010). Predicting and changing behavior: The reasoned action approach. New York: Psychology Press.
- Frey, B. S. and Jegen, R. (2001) 'Motivation crowding theory', *Journal of Economic Surveys*, 15, 5, 589–611.
- Goodin, R. E., Headey, B., Muffels, R. and Dirven, H.-J. (1999) *The Real Worlds of Welfare Capitalism*, Cambridge: Cambridge University Press.
- Groot, N. de en B. van der Klaauw (2019). The effects of reducing the entitlement period to unemployment insurance benefits. In: *Labour Economics*, jg. 57, p. 195-208.
- Groot, L. and van der Veen, R. (2000) 'Clues and leads in the policy debate on basic income in the Netherlands', in R. van der Veen and L. Groot (eds.), *Basic Income on the Agenda: Policy Objectives*

- Groot, L., Muffels, R. J., & Verlaet, T. (2019). Welfare states' social investment strategies and the emergence of Dutch experiments on a minimum income guarantee, *Social Policy & Society* (2019) 18:2, 277–287, Cambridge University Press
- Headey, B., Muffels, R.J.A., & Wagner, G. (2010). Long-running German panel survey shows that personal and economic choices, not just genes, matter for happiness. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)*, 107(42), 17922-17926
- Hemerijck, A. (2013) *Changing Welfare States*, Oxford: Oxford University Press.
- Hemerijck, A. (ed.) (2017) *The Uses of Social Investment*, Oxford: Oxford University Press.
- Hoff Stella Hoff Bart van Hulst Jean Marie Wildeboer Schut Benedikt Goderis (2019), *Armoedekaart 2019* (<http://digitaal.scp.nl/armoedeinkaart2016>), SCP-Den Haag.
- Koning, P. (2012a). Beter een stok dan een wortel. In: J. Ultermark, A.J. Gielen en M. Ham (red.), *Wat werkt nu werkelijk? Politiek en praktijk van sociale interventies* (p. 105-120). Amsterdam: Van Genneep.
- Kluve, J., (2010). The effectiveness of European active labor market programs. *Labour Econ.* 17, 904–918. doi:10.1016/j.labeco.2010.02.004.
- Kok, L., D. Hollanders en J.P. Hop (2006). *Kosten en baten van re-integratie*. Amsterdam: SEO Economisch Onderzoek.
- Koning, J. de, A. Gelderblom, K. Zandvliet en L. van den Boom (2005). *Effectiviteit van re-integratie. De stand van zaken. Literatuuronderzoek*. Den Haag: Ministerie van Sociale Zaken en Werkgelegenheid.
- Koning, J. de, P. de Hek, L. Mallee, F. Rosing en M. Groenewoud (2014). *Uitkomsten en ervaringen experimenten netto-effectiviteit re-integratie*. Rotterdam/Amsterdam: Sero/Regioplan Beleidsonderzoek/Epsilon Research.
- Kremer, M., J. van der Meer, en M. Ham (2017). Van werk naar welbevinden – werkt de zachte hand in de bijstand? *Sociale Vraagstukken* (<https://www.socialevraagstukken.nl/van-werk-naarwelbevinden-werkt-de-zachte-hand-in-de-bijstand/>)
- Mani, A., Mullainathan, S., Shafir, E. and Zhao, J. (2013) 'Poverty Impedes Cognitive Function', *Science*, August, 976–80.
- Marshall, T., Goldberg, R. W., Braude, L., Dougherty, R. H., Daniels, A. S., Sushmita Shoma Ghose, E. D., et al. (2014). Supported employment: Assessing the evidence. *Psychiatric Services*, 65(1), 16–27. <https://doi.org/10.1176/appi.ps.201300262>
- Morel, N., Palier, B. and Palme, J. (2012) *Towards a Social Investment Welfare State? Ideas, Policies and Challenges*, Bristol: The Policy Press.
- Muffels, R. and B. Headey (2013). Capabilities and choices: Do they make Sense for understanding objective and subjective well-being? An empirical test of Sen's capability framework on German and British panel data." *Social Indicators Research*, 110(3): 1159-1185.
- Muffels, R. & S. Vos (2018), *Tussenrapportage No. 1, Vertrouwensexperiment Apeldoorn/Epe*, Tilburg University/Gemeente Apeldoorn/Epe, p. 24
- Muffels, R., K. Stam & S. van Wanrooij (2019), *Tweede Voortgangsverslag Vertrouwensexperiment Apeldoorn/Epe: No. 2, 1 maart 2019*, Tilburg University/ReflecT-Tranzo, p. 1-30.

- Muffels, R. & J. Van der Klink (2017), *Offerte Vertrouwensexperimenten Apeldoorn, Apeldoorn/Epe, Renkum, Epe en Geldrop-Mierlo: Onderzoekdesign en kostenraming onderzoek*, Tilburg University, 10 mei 2017.
- Mullainathan, S. and Shafir, E. (2013) *Scarcity – The True Cost of Not Having Enough*. London: Penguin Books.
- Nazar, Amir (2010), *Wat werkt? Wat weten we over effectiviteit?*, Ministerie van Sociale Zaken en Werkgelegenheid (<http://www.onderzoekwerkeninkomen.nl/rapporten/xbdiusf2>).
- Ryan, R. M., Kuhl, J. and Deci, E. L. (1997) 'Nature and autonomy: an organizational view of social and neurobiological aspects of self-regulation in behavior and development', *Development and Psychopathology*, 9, 4, 701–28.
- Sen, A. K. (2004) 'Capabilities, lists, and public reason: continuing the conversation', *Feminist Economics*, 10, 3, 77–80.
- Sen, A. K. (1999). *Development as Freedom*, New York, Knopf.
- Soentken, M., van Hooren, F. and Rice, D. (2017) 'The impact of social investment reforms on income and activation in the Netherlands', in A. Hemerijck (ed.), *The Uses of Social Investment*, OUP Oxford, Kindle Edition, 475.
- Van Echtelt, Patricia e.a. (2019), *Eindevaluatie van de Participatiewet*, SCP.
- Thaler, R. H. and Sunstein, C. R. (2008) *Nudge: Improving Decisions about Health, Wealth, and Happiness*, New Haven, Conn.: Yale University Press.
- Van der Klink JJ, Bultmann U, Burdorf A, Schaufeli WB, Zijlstra FR, Abma FI, Brouwer S, van der Wilt GJ., Sustainable employability--definition, conceptualization, and implications: A perspective based on the capability approach. *Scand. J Work Environ Health* 2016; 42:71-9.
- Van Ryn, M., & Vinokur, A. D. (1992). How did it work? An examination of the mechanisms through which an intervention for the unemployed promoted job-search behavior. *American Journal of Community Psychology*, 20(5), 577-597.
- Veenhoven, R. (1996). Developments in satisfaction-research. *Social Indicators Research*, 37(1), 821–46. <https://doi.org/10.1007/BF00300268>
- Widerquist, K. (forthcoming) *The Devil's in the Caveats: A Critical Analysis of Basic Income Experiments for Researchers, Policymakers, and Citizens*, Palgrave Macmillan.
- Van Hooft E.A.J. (2016). Self-Regulatory Perspectives in the Theory of Planned Job Search Behavior: Deliberate and Automatic Self-Regulation Strategies to Facilitate Job seeking. In:
- Wennekers, A. J. Boelhouwer, C. van Campen, J. Kullberg (2019), *De sociale staat van Nederland*, SCP Den Haag.

Bijlage A. Resultaten regressieanalyses cliëntvragenlijsten

Tabel A1.1. Niet-gestandaardiseerde regressie coëfficiënten (b) geschat op basis van OLS of logit regressie model (armoede)

Uitkomstmaten ¹	Eigen regie ²	Intensieve begeleiding ²	Meting 2	Meting 3
Welbevinden	0	0	0	.24*
Mentale gezondheid	0	0	0	.37**
Subjectieve gezondheid	0	0	0	0
Keuzemogelijkheden	0	0	0	0
Perceptie sociale participatie	0	0	0	0
Sociale netwerken	0	0	0	0
Vrijwilligerswerk/mantelzorg	0	-3.51*	0	0
Zelfredzaamheid	2.30*** ³	1.63*	0	0
Vertrouwen instituties	0	0	0	0
Sociaal vertrouwen	0	0	0	0
Inkomensarmoede (niet kunnen rondkomen-logitmodel)	-.91*	0	-.63***	-.86***
Deprivatiearmoede (3 uit 5 missen-logit model) ³	-.70*** ⁴	-0.29*** ⁴	0	0

Noot 1: ***p<0,01; **p<0,05; *p<0,10

Noot 1: In dit model waren de uitkomstmaten in de eerste kolom de afhankelijke variabelen Y. De belangrijkste onafhankelijke variabelen waren treatment en meting. Verder is gecontroleerd voor zittend bestand/nieuwe instroom, klantmanagernummer en een aantal controlevariabelen (leeftijd, geslacht, opleiding, duur uitkering, leefvorm, en huisvesting). Deze uitkomsten zijn de test voor de verschillen van de buitenste rij en kolomtotalen van tabellen 3.1.1 t/m 3.1.12. Bij een 0 in de tabel, kon er geen significant effect gevonden worden.

Noot 2: Voor de testen van de effecten van eigen regie en intensieve begeleiding was de standaardgroep de referentiecategorie waartegen werd getest. Bij de effecten van meting 2 en 3 was de nulmeting de referentiecategorie.

Noot 3: Voor de uitkomstmaat zelfredzaamheid vinden we in Epe voor eigen regie een significant positief effect (2.02;p<0.10) . Voorts tonen de schattingsresultaten voor de uitkomstmaat deprivatiearmoede in Epe een negatief effect (minder armoede) aan voor eigen regie (-0.28;p<0.01).

Noot: 4: Het betreft het totale effect van de twee treatments eigen regie en intensieve begeleiding op deprivatiearmoede (minder armoede), berekend als het saldo van het positieve hoofdeffect voor nieuwe instroom en het negatieve interactie effect van beide treatments met het zittende bestand.

Bron: Deelnemers vragenlijsten in de drie metingen Vertrouwensexperiment Apeldoorn/Epe, Tilburg University, 2020.

Tabel A1.2: Regressie uitkomsten interactiemodel treatment en meting productevaluatie (significante treatment effecten per meting)

Uitkomstmaten	Eigen regie			Intensieve begeleiding			Standaardgroep		
	Nul-meting	M2	M3	Nulmeting	M2	M3	Nul-meting	M2	M3
Welbevinden	0	0	0	0	0	0	0	0	0
Mentale gezondheid	0	0	0	0	0	0	0	0	0.48*
Subjectieve gezondheid	0	0	0	0	0	0	0	0	0
Keuzemogelijkheden	0	0	0	0	0	0	0	0	0
Perceptie sociale participatie	0	0	0	0	0	0	0	0	0
Sociale participatie	0	0	0	0	0	0	0	0	0
Vrijwilligerswerk	0	0	0	0	0	0	0	0	0
Zelfredzaamheid	1,61**	1.55*	1.54*	0	0	0	0	0	0
Vertrouwen instituties	0	0	0	0	0	0	0	0	0
Sociaal vertrouwen	0	0	0	0	0	0	0	0	0
Inkomensarmoede	-.59*	0	0	0	0	0	0	0	0
Deprivatiearmoede (3 uit 5)	0	0	0	0	0	0	0	0	0

Noot 1: *** $p < 0,01$; ** $p < 0,05$; * $p < 0,10$

Noot 2: voor de effecten van de standaardgroep in meting 2 en 3 is de standaardgroep in de nulmeting de referentiecategorie. Voor de effecten van eigen regie in de nulmeting is de referentie de standaardgroep in de nulmeting en voor metingen 2 en 3 geeft de parameter het effect weer van de groep in die meting dat bovenop het effect komt van het effect van de eigen regiegroep in de nulmeting waar de standaardgroep de referentie is. Hetzelfde geldt voor het effect van intensieve begeleiding.

Noot 3: In dit model zijn de uitkomstmaten in de eerste kolom de afhankelijke variabelen Y. De belangrijkste onafhankelijke variabelen zijn treatment en meting en de interactie tussen deze twee. Verder is gecontroleerd voor zittend bestand/nieuwe instroom, klantmanager/consulent nummer en een aantal controlevariabelen (leeftijd, geslacht, opleiding, duur uitkering, leefvorm, en huisvesting). Deze uitkomsten zijn de test voor de verschillen van de tussenliggende cellen van tabellen 3.1.1 t/m 3.1.12. Bij een 0 in de tabel, is er geen significant effect gevonden.

Bijlage B. Resultaten regressieanalyses procesevaluatievragenlijsten: uitkomstmaten

BIJLAGE B

Tabel 1: Niet-gestandaardiseerde regressie coëfficiënten (*b*) geschat op basis van OLS regressiemodel per treatment groep over de tijd en per meting

Uitkomstmaten	Eigen regie	Intensief	Meting 2	Meting 3
Motivatatie eigen treatment	-3,1**	0	0,6***	0,7**
Motivatatie hoogst bereikbare doel	-1,9*	0	1,3***	1,4***
Realistische eisen	0	0	1,0***	1,0***
Zelf in actie komen	-2,2**	0	1,5***	1,6***
Zelf doelen stellen	0	0	1,1***	1,3***
Verwacht effect treatment	-3,0**	0	-0,5***	-0,3*
Sociale participatie	-1,7**	0	0,8***	0,9***
Participatieladder	-1,4**	0	1,0***	1,0***
Leereffect klant	n.s.	0	-0,3**	-0,4**
Leereffect klantmanager	n.s.	0	1,2***	1,3***

Noot 1: *** $p < 0,01$; ** $p < 0,05$; * $p < 0,10$

Noot 2: Voor het testen van de effecten van eigen regie en intensieve begeleiding is de standaardgroep de referentiecategorie. Bij de effecten van meting 2 en 3 is de standaardgroep in de nulmeting de referentiecategorie.

Noot 3: In dit model zijn de concepten in de eerste kolom de afhankelijke variabelen *Y*. De belangrijkste onafhankelijke variabelen zijn treatment en meting. Verder is gecontroleerd voor zittend bestand/nieuwe instroom en klantmanagernummer. Deze uitkomsten geven de test weer voor de verschillen van de buitenste rij en kolomtotalen van tabellen 3.3.1 t/m 3.3.10. Bij n.s. (niet significant) in de tabel, is er geen significant effect gevonden.

Bron: Vragenlijsten procesevaluatie Vertrouwensexperiment Apeldoorn, Tilburg University, 2020.

Tabel 2: Niet-gestandaardiseerde regressie coëfficiënten (b) geschat op basis van OLS interactie model (treatment effecten per meting)

Standaardgroep			Eigen regie			Intensief		
Uitkomstmaat	M2	M3	Nulmeting	M2	M3	Nulmeting	M2	M3
Motivatatie eigen treatment	1,5***	1,2***	-2,5**	-1,5***	-0,8*	0,8*	-1,4***	-0,8*
Motivatatie hoogst bereikbare doel	1,7***	1,4***	0	-0,7*	0	0	-0,7*	0
Realistische eisen	1,8***	1,5***	0	-1,5***	0	1,1**	-1,1**	-1,2**
Zelf in actie komen	2,3***	1,9***	0	-1,5**	0	0	-1,1**	0
Zelf doelen stellen	1,8***	1,6***	0	-1,1**	0	1,2**	-1,2**	-1,1**
Verwacht effect treatment	0	0,6*	-2,5**	-1,3***	0	0	-0,8**	-1,8***
Sociale participatie	1,2***	1,4***	-1,4*	0	0	0	-1,0**	-1,0**
Participatieladder	1,2***	1,4***	-1,3**	0	-0,9***	0	-0,4**	-0,6**
Leereffect deelnemer	0,4**	0	0	-1,4***	0	0	-0,9**	-1,5***
Leereffect consulent	1,9***	2,3***	0	-0,9**	-2,1***	1,1**	-1,5***	-1,2**

Noot 1: ***p<0,01; **p<0,05; *p<0,10

Noot 2: voor de effecten van de standaardgroep in meting 2 en 3 is de standaardgroep in de nulmeting de referentie categorie. Voor de effecten van eigen regie in de nulmeting is de referentie de standaardgroep in de nulmeting en voor de vervolgmetingen geeft de parameter het effect weer van de groep in die meting dat bovenop het effect komt van de eigen regiegroep in de nulmeting waar de standaardgroep de referentie is. Hetzelfde geldt voor het effect van intensieve begeleiding.

Noot 3: In dit model waren de concepten in de eerste kolom de afhankelijke variabelen Y. De belangrijkste onafhankelijke variabelen waren treatment en meting en de interactie tussen deze twee. Verder is gecontroleerd voor zittend bestand/nieuwe instroom en klantmanager/consulent nummer. Deze uitkomsten zijn de test voor de verschillen van de tussenliggende cellen van tabellen 3.2.1 t/m 3.2.10. Bij n.s. (niet significant) in de tabel, kon er geen significant effect gevonden worden.

Bron: Vragenlijsten nulmeting en vervolgmetingen procesevaluatie Vertrouwensexperiment Apeldoorn, Tilburg University, 2020.